Four Essays in Macroeconomics

Inauguraldissertation zur Erlangung der Würde eines Doctor rerum oeconomicarum der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Universität Bern

vorgelegt von
Ronald Indergand
von Silenen, Uri

2016

Die Fakultät hat diese Arbeit am 10. November 2016 auf Antrag der beiden Gutachter Prof. Dr. Klaus Neusser und Prof. Dr. Ulrich Woitek als Dissertation angenommen, ohne damit zu den darin ausgesprochenen Auffassungen Stellung nehmen zu wollen.

Acknowledgement

Without my mentors, senior colleagues and friends, I would not have been in the position to complete this dissertation.

First and foremost, I thank Prof. Klaus Neusser for his efforts both as a teacher during my bachelor and master studies, and as a supervisor of my PhD thesis. I am extremely grateful for the flexibility he granted me with respect to the choice of the topics and for his unconditional support during my successful applications for research grants and fellowships. I am also very grateful to my co-supervisor, Prof. Ulrich Woitek, who supported me unconditionally during all stages of my PhD. My sincere thanks go to my co-authors, Dr. Andreas Beerli and Dr. Stefan Leist, for being a source of inspiration, help and motivation during our joint projects.

I also thank the State Secretariat for Economic Affairs (SECO) for supporting several of the chapters contained in this PhD thesis. Most importantly, Dr. Bruno Parnisari made it possible for me to work part-time at the SECO and devote the rest of my time to research. He further allowed my co-author Dr. Stefan Leist and me to collect and publish a real-time data set for Switzerland. I am also very grateful to Dr. Christoph Sax, my former colleague at the SECO, who introduced me to the R programming language, offered his support for many related questions and has become a dear friend.

For their financial support for research stays abroad, I thank the Berne University Research Foundation, the Irène und Max Gsell Stiftung and the Swiss National Science Foundation. Particular thanks go to Dr. Tom Stark and the Real-Time Data Research Center at the Federal Reserve Bank of Philadelphia, to Prof. Francis X. Diebold (University of Pennsylvania) and to Prof. David Romer (University of California, Berkeley) for their hospitality and support during my research visits in the United States.

Dr. Arnd Klein, Dr. Jan-Aaron Klaassen and Thomas Wüthrich have not only become very close friends during graduate courses but were always great partners for intriguing discussions concerning both academic and non-academic questions. For being great mentors and friends, I thank Geli Spescha and Dr. Werner Aeberhardt, who both offered me support and advice whenever I needed it. I am extremely grateful to my parents, Hubert Indergand and Barbara Indergand, as well as my grandmother, Anna Meisslitzer, for their encouragement and unconditional support during every stage of this dissertation and my life. Their interest in and discussions about society, politics and economics sparked my early curiosity about these topics and made me choose economics as the center of my studies and professional work.

Finally, my affectionate thank goes to Isabel Martínez for her patience and her advice. I am thankful to Isabel that she encouraged me to set out to the United States with her and spend a year abroad which not only helped me for my research but allowed me to develop as a person. Most of all, I thank her for her love and her support during all the ups and downs I experienced on this journey.

Bern, August 10 2016

Ronald Indergand

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Without reliable data, most applied economic research would not be possible. Fiscal policy would have to manage without quantitative benchmarks and monetary policy would be left with little guidance. Accordingly, virtually all countries invest considerable resources into the collection, the preparation and the timely provision of economic data. However, the compilation of such data is often not straightforward. A few variables can be comparatively easily measured, for example the prices of (liquid) stocks or government bond yields. Other variables, such as Gross Domestic Product (GDP), are not only based on abstract, complicated and sometimes changeing concepts but can never be directly observed and measured.

Three forms of uncertainty arise from this. First, there are measurement problems arising during the survey, projection and estimation processes. Sampling errors in surveys or statistical errors in regression analyses are typical examples. Often, more information becomes available over time and statistical agencies therefore revise their estimates frequently. Hence, the estimates often become more precise over time and the agencies are faced with a trade-off between timeliness and accuracy (see also Manski, 2015, who distinguishes between transitory, permanent and conceptual uncertainty).

Second, there exists uncertainty about the measurement concepts. Eventough national accounting was set on a sound and clear foundation during the middle of the last century, in particular through work by Richard Stone which won him the Nobel Price in 1984, definitions and conventions for measuring economic output are subject to change. This means that the concept (definition) of a variable, as measured in the 1970s, may not correspond exactly to the same concept when measured today. For example, software purchases by firms were for a long time not counted as an investment but as an intermediate input whereas today it forms part of investment in equipment and thus GDP.

Third, there exist different measurement techniques and they are continuously evolving. For example, national accounting on the quarterly frequency may be based either on representative surveys, such as in the US, or alternatively on temporal disaggregation of annual data, such as in Switzerland. Good practice in both fields may change over time and new approaches are developed as time passes. Further examples include the different techniques that extist for imputing missing data, such as carrying forward the last observation, ARIMA forecast or regression imputation.

Even for variables that are based on simple concepts and that are relatively straightforward to collect, the data provider may face big challenges. For example, Swiss unemployed have to register in one of about 130 local job mediation agencies to get unemployment benefits. To compute the total of all registered unemployed, a federal statistical agency simply adds up the figures of all local agencies.[1] The data which gets the attention, however, is the seasonally adjusted (SA) version of this total. To compute the SA data, the data provider has to employ ARIMA modelling and estimation strategies as well as signal extraction techniques such as the X-11 algorithm, Wiener-Kolmogorov filtering, or state space modelling and Kalman filtering. In this process, the statistical agency has to make many important decisions about the modelling of the series and the extraction of the unobserved SA data. All these decisions have consequences on the outcome. In fact, they may completely alter the final, seasonally adjusted series, particularly for the most recent data points which get the largest attention in general.

[1] If one is interested in the actual unemployment rate (based on internationally harmonized concepts), one has to consider that many individuals without job but looking for work are not registered in one of these local job mediation agencies. To include these individuals, one has to consult other data sources, such as labour force surveys.

Due to all these reasons, the measurement of economic variables is a very challenging issue. Statistical agencies are forced to revise their data frequently and sometimes heavily even in case of the most important economic variables, such as labour market data, monetary aggregates, or GDP. In recent years, revisions to GDP figures have regularly attracted the attention of the broad public audience (Plickert, 2014; Müller, 2015; Economist, 2016) as well as scholarly research. However, that economic data may change with new data releases is not a new insight and was already well-known in the early developments of the modern national accounting principles (see for example Kuznets, 1937 and Stone, Champernowne and Meade, 1942). After national accounting has been established as a regular task of national statistical agencies in western countries, several studies have tried to quantify the size of data revisions to GDP and other variables (starting with Morgenstern, 1950 and Zellner, 1958) and tested them for biases (see Glejser and Schavey, 1974 for an early example). Soon it became clear that data revisions have a large influence on economic forecasting (Howrey, 1978).

Until recently, however, almost all authors had to build their own real-time data set in order to analyze the effect of data revisions. Such a data set is a systematic collection of data vintages, that is, older versions of a data series as they were released by the statistical agencies at different points in time. This process can be very cumbersome and often involves archival work if the data has not been stored in a systematic way. In that case the vintages have to be collected from old press releases or printed statistical bulletins.

The database compiled by Croushore and Stark (2001) is the first comprehensive real-time data set for the US which has been made publicly accessible. It has been extended considerably and is updated regularly until today. Other real-time databases have followed for many countries and variables, offering a rich data source for economic research. The amount of studies that was initiated from this during recent years is impressive (see the literature reviews by Croushore, 2011*b* and Cimadomo, 2016). A large part of this dissertation follows in this tradition with a focus on Switzerland.

As in any other country, research employing real-time data in Switzerland can only advance significantly if an appropriate data set and accompanying meta information about the release dates, the revision policies, the data construction process and alike are readily available. The first contribution of this PhD thesis is the compilation and publication of such a data set for Switzerland, consisting of about 20 macroeconomic variables (Chapter 2, joint with Stefan Leist). The data set is maintained and updated by the State Secretariat for Economic Affairs (SECO) and can be downloaded from its website.[2]

Based on this data set, a thorough analysis of revisions to Swiss national accounts data follows in Chapter 3. More specifically, all data revisions after the publication of the preliminary estimate should be unpredictable. It may be reassuring for both, the data suppliers and its users that this is indeed the case for Swiss GDP. However, for all expenditure-side components of GDP, the most important finding of my research is that data revisions reduce extreme growth rates. This means that preliminary announcements tend to be corrected towards their mean in later releases. This result may be surprising given that the preliminary Quarterly National Accounts (QNA) data in Switzerland is calculated based on best (minimum variance) linear unbiased extrapolation techniques (Chow and Lin, 1971). However, the finding is inline with several studies focusing on GDP of other countries (Faust, Rogers and Wright, 2005) as well as other variables including monetary aggregates in the US (Mankiw, Runkle and Shapiro, 1984; Kavajecz and Collins, 1995) or monthly data on employment, unemployment and industrial production in the UK (Garratt and Vahey, 2006).

Chapter 4 is an attempt to explain the finding that mean reversion is such a predominant feature of macroeconomic data. I study seasonal adjustment as a specific but very important source for data revisions. Based on three different approaches, namely using Swiss unemployment and employment data as real-world examples, running simulations using GDP data, and running simulations based on artificial data, I show that seasonal adjustment (in particular X-11 based) methods very often produce preliminary data with a mean-reverting revision process. While the exact source for this tendency

[2] Updated versions can also be obtained from the author of this dissertation.

remains unclear, this provides a starting point for future research that attempts to mitigate forecast inefficiency in seasonally adjusted economic data.

Chapter 5 (joint with Andreas Beerli) turns to another, policy relevant question in the field of macro labor economics. In recent decades, the skill-mix of migrants has undergone significant changes. Migrants in developed countries have become more highly skilled and more positively selected. In Switzerland, the share of new immigrants with a tertiary education almost tripled from 17% to 47% between 1990 and 2010, surpassing the educational upgrading of the native population by far. While the public discussion has often focused on the role of the more liberal migration policy (free movement of persons treaty of Switzerland with the EU), we argue that much more fundamental forces are at play, reshaping the structure of labor markets in advanced economies more generally.

By combining two different strands of the literature we show that routinization as a form of skill-biased technical change explains to a large extend what we are observing in the skill-mix of migrants. More specifically, we rely on a model for the selection and geographical sorting of immigrants by Grogger and Hanson (2011) and on new advances in the literature of skill-biased technical change by Autor and Dorn (2013) explaining job market polarization by routinization, that is, the substitution of routine-intensive labour by computer capital.

We establish that the labour market in Switzerland, like in many other developed countries, is characterized by growth at the poles of the skill and wage distribution (job market polarization). Both the number of workers and their wages grow at the top and (more mildly) at the bottom of the income distribution, whereas a large contraction can be seen in middle-paying, middle-skill jobs.[3] We then use a measure of local, technology-driven demand-shifts (Autor and Dorn, 2013) to identify the effect of demand-pull forces and separate them from the influence of changes in educational attainment in origin countries, a supply-push factor, as well as changes in immigration restrictions.

[3] Note that this does not necessarily mean that "the middle class disappears" (in fact, the former middle class could grow under such a scenario). However, it does imply that we see shifts in the earings distribution.

We find that the skill composition of immigrants in Switzerland is strongly demand-driven. The introduction of automation technologies such as computers or industrial robots has replaced a large share of clerks and blue-collar workers in manufacturing and construction since the 1980s. On the other hand, these new technologies have raised the demand for highly skilled professionals such as managers, engineers or scientists, who complement the use of automation technologies. The demand for low skilled workers employed in service-oriented manual tasks, such as social workers or hair dressers, has been less affected as they are less prone to substitution by computer capital. Accordingly, while migrants predominantly worked in blue-collar occupations prior to the 1980s, they show much stronger attachment to high-skill jobs in recent years. The change of educational supply in the origin countries would predict the strongest inflow of workers at middle skill levels. Yet, far stronger gains actually occurred at the top (tertiary education) with very modest gains below, which highlights the important role of skill demand in the destinations. Finally, immigration policy may qualify these effects to some degree, depending how it affects the incentives to migrate for workers with different skills. In the Swiss case, however, the policy changes of the recent past had a very mild, even adverse, effect on the skill composition of immigrants.

Taken together, all four chapters contribute to important macroeconomic questions in Switzerland. The first three chapters provide new evidence on the accuracy of Swiss national accounts and other data. By offering both a new data set and an explanation for mean reversion in preliminary data releases, this work may help statistical agencies in their permanent work to improve the quality of Swiss macroeconomic data. At the same time, this work is a base for and points to exciting future research on real-time data and seasonal adjustment methods. The fourth chapter provides new insights on fundamental changes in the Swiss labour market and their effect on migrants as well as the role of migration policy. In light of the recent policy discussion including the referendum about the free movement of persons treaty, I hope that this work helps to clarify some of the issues and to keep the often controversy debates objective.

Chapter 2

A Real-Time Data Set for Switzerland

Ronald Indergand and Stefan Leist

Summary: Accessibility of real-time data is crucial for applied macroeconomic researchers who aim at evaluating forecasts, policy decisions or the accuracy of initial data estimates. To the extent of our knowledge, no appropriate and comprehensive real-time data set has been published for Switzerland so far. This paper introduces such a data set, which can be downloaded online. The balanced database includes quarterly, seasonally adjusted vintages of the most important economic variables on the national level. A short analysis of data revisions is provided for quarterly GDP. The magnitude of revisions are comparable to other countries such as the Euro Area or the United States. However, revision policy may differ considerably and potentially influence a statistical analysis.

This project was supported by the State Secretariat for Economic Affairs, Switzerland (SECO). The views and opinions expressed in this paper are solely those of the authors and are not necessarily those of the State Secretariat for Economic Affairs. Particular thanks go to Bernhard Weber (SECO) who kindly provided us with historical vintages of Swiss employment. We are also very grateful to Isabel Martinez, Prof. Klaus Neusser, Bruno Parnisari, Christoph Sax and Peter Steiner who supported us during several stages of the project.

2.1 Introduction

How informative are initial estimates of macroeconomic data? How good is a specific forecasting model? Are empirical findings robust to data revisions? A real-time data set is needed to address these questions. The literature examining the relevance of data-revisions for applied economic research acknowledges this necessity and has a long tradition. Moreover, it is expanding rapidly after comprehensive data sets have become available for many countries (see below). In what follows, we provide a short overview of typical applications of such data. See Croushore (2011*b*) or an extensive review of the literature.[1]

Initiated by the seminal work of Diebold and Rudebusch (1991), a canonical view concerning forecasting excercises has emerged which asserts that, in the presence of data revisions, any evaluation of economic predictions should rely on real-time data sets (see, e.g., Croushore, 2011*b*). Otherwise, forecasters might rely on a model which provides reasonable forecasts in a pseudo-out-of-sample[2] exercise but potentially contains indicators that are heavily revised and, therefore, probably not useful in a forecasting context. As a result, it has become standard in the literature to work with real-time data for assessing the quality of forecasting models.

Similarly, if the aim is to evaluate policy decisions of the past, ignorance of revisions may lead to erroneous conclusions (see, e.g., Orphanides, 2001, Croushore and Evans, 2006). Moreover, Orphanides and Van Norden (2005) show that real-time issues are important for the assessment of inflationary pressures. Furthermore, the results of empirical economic research in general may depend on the data vintage at hand (Anderson, 2006). Indeed, Croushore and Stark (2003) show that some key macroeconomic results pre-

[1] Dean Croushores bibliography on real-time data literature also contains a large collection of papers: `https://facultystaff.richmond.edu/~dcrousho/`

[2] In a pseudo-out-of-sample forecasting exercise data ranging from $t = 1, \ldots, T$ is available. To perform forecasts you then estimate a model using only t=1,...,T-k observations, where k is an arbitrary chosen integer. The forecasts from this model generate data for $t = T - k + 1, \ldots, T$ which is compared to the actual value. This approach provides information on the preciseness of forecasts only if the data is not subject to revisions.

sented in the literature may not be robust to the selection of different data vintages.

Another strand of literature examines the efficiency of first data releases by the statistical agencies. A predictable revision process would imply potential for improvement of first releases. Faust, Rogers and Wright (2005), Aruoba (2008) and Corradi, Fernandez and Swanson (2009) address this question and find predictable revisions to GDP and other variables for several G7 countries. Other authors model the data revision process in order to extract information about the true values and improve on the preliminary data (see, e.g., Jacobs and Van Norden, 2011, Cunningham et al., 2012).

Applied research employing real-time data for Switzerland is rather scarce to date. Some notable exceptions include Amstad and Fischer (2009) who are the first to show that weekly updates of inflation-nowcasts are informative. Jordan et al. (2005) and Cuche-Curti, Hall and Zanetti (2008) show that revisions to Swiss GDP are an important source of uncertainty for monetary authorities and may lead to suboptimal monetary policy decisions. Kholodilin and Siliverstovs (2012) employ a partly real-time data set to show that business tendency surveys (BTS) can be useful for predicting GDP using a dynamic factor model. Other authors find that BTS are able to predict revisions in the case of GDP during a short subperiod of their total sample (Siliverstovs, 2011), as well as in the case of current account data for certain types of revisions (Jacobs and Sturm, 2008). Finally, Siliverstovs (2013) shows that the chronology of the Swiss business cycle is robust across real-time data vintages.

The above mentioned research has been spurred by the release of real-time databases. The most notable examples of such databases include Croushore and Stark (2001) for the US[3], Giannone et al. (2012) for the Euro-Area[4], Egginton, Pick and Vahey (2002) and Castle and Ellis (2002) for the UK[5]

[3] `http://www.phil.frb.org/research-and-data/real-time-center/real-time-data/`

[4] `http://www.eabcn.org/eabcn-real-time-database`

[5] `http://www.econ.cam.ac.uk/research/keepitreal/keepitreal/index.htm` and `http://www.bankofengland.co.uk/statistics/Pages/gdpdatabase/default.aspx`

and Knetsch (2009) for Germany[6]. Furthermore, there exists a real-time database provided by the OECD[7] including many variables for most of its member states. Fernandez et al. (2012) extend this data set by short windows of historical vintages (consisting of about 10 quarters for each vintage).[8]

So far, to the best of our knowledge, Swiss researchers have not been given access to a public and comprehensive real-time data set.[9] The OECD real-time database includes data for Swiss variables. However, the data is heavily rounded in case of several vintages and variables (e.g., GDP), yielding zero-growth rates for many quarters. Additionally, the usage of the data proves to be somewhat cumbersome as the edition dates (vintage dates) do not always correspond to the official release dates of the QNA figures.

We fill this gap for Switzerland by providing real-time data on national accounts, labor market statistics, prices and interest rates. For comparison and because measures for foreign economic activity are vital in models of a small open economy, we also include real-time data of US, Euro Area 17 (EA 17) and Japanese GDP. The construction of the data set is described in Section 2. Section 3 provides a brief analysis of revisions to Swiss GDP and compares it to other countries. We find that revisions in Switzerland are similar in scope as in other countries. However, country-specific particularities of the revision processes may influence a statistical analysis and should be taken into account in international comparisons.

2.2 The Data Set

2.2.1 Definitions

Real-time data for a given variable is stored in matrix form, as Table 2.1 shows. The row labels represent the usual time axis, t, of a time series whereas the column labels consist of the different release dates, v, of the series. Such a column is commonly referred to as a vintage and contains the

[6] `http://www.bundesbank.de/Navigation/EN/Statistics/Time_series_databases/Real_Time_Data/realtime_zeitreihen_node.html`
[7] `http://stats.oecd.org/mei/default.asp?rev=1`
[8] `http://www.rthd-oecd.org`
[9] Bernhard (2016) will provide vintages for Swiss GDP.

Table 2.1: Vintages of Swiss GDP

	2006 Q1	2006 Q2	2006 Q3	2006 Q4	2007 Q1	2007 Q2	2007 Q3
2004 Q1	107035786100	106958170500	107130657300	107123282400	107108887400	107100207600	109299794900
2004 Q2	107084259100	107065443100	107454478300	107448934800	107452424500	107446665600	109726416900
2004 Q3	107220869400	107229818500	107614244100	107628360700	107627164900	107631258200	109625977000
2004 Q4	107301969300	107407358100	107816675100	107818905900	107833411200	107855056800	110195541200
2005 Q1	107815565000	107724200300	108170675400	108159942800	108141632100	108117321900	110911475100
2005 Q2	108773313000	108738147600	109109336600	109095140300	109100042700	109088799000	111893952100
2005 Q3	109700609400	109701542700	110105684900	110132200400	110127193700	110134642800	112783236500
2005 Q4	110277485500	110424785000	110876039000	110878107000	110893293200	110934991700	113838314600
2006 Q1	NA	111457016200	111709154200	111648104600	111637322000	111594859300	114804479600
2006 Q2	NA	NA	112527759600	112327045500	112386743400	112367559800	115530816900
2006 Q3	NA	NA	NA	112739072000	112852772700	112862703500	116501235800
2006 Q4	NA	NA	NA	NA	113364145400	113429946600	117127892200
2007 Q1	NA	NA	NA	NA	NA	114316995400	117896170800
2007 Q2	NA	NA	NA	NA	NA	NA	118769102400
2007 Q3	NA	NA	NA	NA	NA	NA	NA
2007 Q4	NA	NA	NA	NA	NA	NA	NA

Note: This is a screenshot of the printed vintages in the R console (RStudio). The data in the excel-files also downloadable are organized in the same manner.

most up-to-date information available at that time for the series. In what follows, we denote the time index t as a subscript and the vintage time index v as a superscript, e.g., x_t^v. The natural log of a variable X is denoted by x.

Our data set consists of vintages as they were available at a particular release date of the quarterly national accounts (QNA). We chose the QNA release date, rather than another point in time, for the following reason. Shortly after the QNA data are published, various professional forecasters update their economic predictions using the most recent information available. Thus, for any forecast evaluation excercise - possibly involving the professionals' forecasts as benchmarks - it is important to have their state of information readily available. The QNA estimates for Switzerland are computed by the State Secretariat for Economic Affairs (SECO) about 60 days after the end of a quarter. For example, $GDP^{2014q2}_{1980q1-2014q1}$ was published on May 28, 2014 (i.e., the release quarter is $v = 2014q2$) and contains quarterly data on GDP starting from the first quarter 1980 to the first quarter 2014.[10]

[10] The release dates varied somewhat but usually were in March, at the end of May / beginning of June, in September and at the end of November / beginning of December. We relied on a quarterly notation in the superscript to enhance readability. The exact publication dates are included in the database whenever possible.

2.2.2 The Variables

Table 2.2 provides an overview of the variables covered by the database so far. The data set is balanced in order to facilitate the usage of the data, i.e., the vintages were retropolated or completed when necessary. We performed seasonal adjustment if seasonal patterns have not already been removed beforehand. Any of these adjustments and transformations are outlined below and are flagged in the data set. Hence, if desired, the user can choose to work with the unbalanced data. The seasonally unadjusted vintages are also provided for the variables that were seasonally adjusted by the autors.

The first vintage is denoted by $v = 2002q4$ and contains data ranging from $1980q1$ to $2002q3$. All subsequent vintages also start in $1980q1$ and each vintage contains an observation more than the vintage before as well as all revised data. The final vintage is $v = 2014q2$ containing data from $1980q1$ to $2014q1$. The data set can be downloaded from this journal's homepage (`http://www.sjes.ch/published.php?Year=2014`), including some metadata. All subsequent updates, changes or extensions of the data set will be documented and provided on `http://www.seco.admin.ch/themen/00374/00456/`.

Table 2.2: Variables Contained in the Data Set

National Accounts Data	GDP (real, sa)
	Government Consumption (real, sa)
	Private Consumption (real, sa)
	Gross Investment (real, sa)
	Total Exports (real, sa)
	Total Imports (real, sa)
	GDP Deflator (sa)
Prices & Financials	Consumer Price Index (nsa, sa)
	Nominal Interest Rate
	Real Interest Rate (ex-post)
Labor Market Data	Employment (nsa, sa)
	Registered Unemployed (nsa, sa)
	Unemployment Rate (nsa, sa)
Foreign Data	Foreign GDP Proxy (real index, sa)
	GDP, USA (real, sa)
	GDP, Euro Area 17 (real, sa)
	GDP, Japan (real, sa)

National Accounts

For GDP, government and private consumption, investment, exports, imports and the GDP deflator, the main sources were the archives of the SECO and the OECD real-time database, provided the data in the latter was not rounded too heavily. All variables were available seasonally adjusted and, apart from the GDP deflator, in volume terms (usually chain-linked[11]). For some variables, j, the start of the time index t varied across vintages. In these cases, we used growth rates of the previous vintage to retropolate the data in order to achieve a balanced data set, starting in $1980q1$ $\forall v, j$. Table 2.A.1 provides the details.

Labor Market

Vintages for total employment (full time equivalents) are available for $v = 2002q4 - 2014q2$.[12] For $v \in \{2011q2, 2011q3\}$, the Swiss Federal Statistical Office (SFSO) released employment with a delay - that is, after the QNA estimate took place. In these cases, the SECO relied on internal estimates which we included in order to achieve a balanced data set. All vintages of total employment were seasonally adjusted using the automated procedure of X-13ARIMA-SEATS without outliers.[13]

The unemployment rate was constructed with the registered unemployed in the nominator and the labour force in the denominator. In principle, we followed the official calculation, where the denominator, i.e., the working population (WP) from the Census, remains fixed for about a decade. It gets updated approximately three years after the new Census has been conducted

[11] The transition from volumes measured at constant prices to the chain-linking method occured by the end of 2003 for the yearly series. The QNA got adjusted to these yearly aggregates in $2004q1$ but experienced further revisions in the course of 2004 and 2005 (see Section 2.31).

[12] For $v = 2002q4 - 2012q3$ they were provided by Bernhard Weber (SECO).

[13] Although this is a rather ad-hoc approach, we believe that it delivers useful seasonally adjusted data. Appendix 4 offers some details on the models chosen for seasonal adjustment. However, the real-time data is also provided not seasonally adjusted. Hence, researchers are enabled to perform there own seasonal adjustement.

and then remains fixed for about another decade.[14] However, this procedure is agnostic of any quarterly changes in the total number of employees and hence, in case of a growing workforce, somewhat overestimating the unemployment rate as time passes. In order to dampen this effect, we linearly interpolated the growth of the workforce between two Censuses and only held the most recent figure from the Census constant as formally described in Appendix 2. Since this adjustment of the denominator is rather ad-hoc, we also included vintages of the registered unemployed (no revisions) in the data set. Seasonal adjustment of the resulting unemployment rate and of the registered unemployed was conducted using the automated procedure of X-13ARIMA-SEATS without outliers. Appendix 4 offers some details.

Prices and Interest Rates

Vintages for the consumer price index (CPI) were taken from the OECD real-time database and from the SFSO homepage. The same procedure for seasonal adjustment as above was applied to all vintages of the CPI. An ex-post real interest rate was computed by substracting year-on-year CPI-inflation (nsa) from the nominal interest rate, which is not subject to revisions.

$$RR_t^v = I_t - 100 \cdot (cpi_t^v - cpi_{t-4}^v) \tag{2.1}$$

The 3-month London Interbank Offered Rate (LIBOR)[15] is used as a nominal interest rate, I_t, expressed in percentage points. The British Bankers' Association started the official calculation in 1989m1. In order to achieve a balanced data set, we retropolated the LIBOR series using monthly growth rates of the 3-month Swiss Franc Euromarket interest rate.[16] The correlation between the two rates during the overlapping period (1989m1 to 2007m10) is 0.998.

[14] See, e.g., SECO (2014*a*). Starting in 2010, annual register data supplemented by surveys serves as a new Census and replaces the former decennial Census. Therefore, the base may be changed more frequently in the future.

[15] Downloadable from `http://www.snb.ch/de/iabout/stat/statpub/statmon/stats/statmon/statmon_E1`

[16] Downloadable from `http://www.snb.ch/de/iabout/stat/statpub/histz/id/statpub_histz_actual`

Foreign GDP

Vintages of real, seasonally adjusted GDP for the US, Japan and the EA 17 were taken from the OECD real-time database. In case of the EA 17, the OECD database does not provide any information for GDP_{2008q1}^{2008q2} and $GDP_{2008q1-2008q2}^{2008q3}$. However, quarterly estimates for the E.A. 15 were publicly available and growth rates can be accessed in Eurostat (2008*a*,*b*). We used this data to extrapolate the missing GDP figures for the EA 17.
In order to obtain a proxy for foreign economic activity (GDP^f), we weight the growth rates of real, seasonally adjusted GDP of the three regions by their share of the sum of nominal GDP ($NGDP$) measured in Euros.[17] Because the nominal levels of the most recent vintage for the euro area were often published with a delay, we use the vintage published one quarter earlier ($v-1$), see Appendix 3 for a detailed exposition. To convert US and Japanese $NGDP$ to Euros we use the spot exchange rates.

Using these growth rates for foreign economic activity, we construct an index for a proxy of the foreign GDP levels starting with 1980q1 = 100.

In the remainder of this paper, we provide a short revision analysis for Swiss GDP. It serves as an example for a possible application of the database and illustrates that revisions to national accounts data in Switzerland may be large and similar to those in other countries.

2.3 An Example: Revisions to Swiss GDP

2.3.1 Why Revisions?

There are several causes for revisions to a published data series. In case of the Swiss QNA data, we classify the reasons for revisions into five groups.[18]

[17] An alternative to consider would be a trade-weighted proxy, requiring real-time vintages of exports by countries, however.
[18] The Swiss QNA benchmark revisions correspond to the definition of major revisions in Eurostat (2013). The other four types of Swiss QNA revisions may all be classified as routine revisions.

Revisions to Quarterly Indicators

The QNA estimates in Switzerland are based on the inter- and extrapolation techniques described and implemented in R by Sax and Steiner (2013). Thus, every revision to an indicator used for temporal disaggregation affects the final quarterly series. Revisions to the quarterly indicators may occur at any time and usually affect the most recent quarters of an indicator.

Revisions to Annual Base Data

The aggregates of the annual national accounts (ANA) are computed by the SFSO and get released every year in August. At the same time the two previous years get revised. For example, in August 2013, the SFSO released annual figures for 2012. In addition, the ANA data for 2011 got revised for the first time and the figures for 2010 for the second time. Yearly data before that is regarded as definitive and remains unrevised apart from benchmark revisions. Whenever the SFSO releases new ANA data, the QNA data computed by the SECO are to be adjusted to those yearly reference values.

Changes in Methodology (Benchmark Revisions)

Concepts as well as the methodology for the computation of national accounts are subject to development. For example, Swiss national accounts may be adjusted to new international standards, or the method of computing quarterly GDP may fundamentally alter. Such changes are large and may affect the dynamics and levels of the whole time series. Usually, they are not introduced gradually but at a specific date and are often called benchmark revisions. The vintage time span covered by our data set, for example, in case of Swiss quarterly GDP includes five benchmark revisions. In $v = 2004q1$ all QNA variables were preliminarily adjusted to the new yearly aggregates which got revised according to the European System of Accounts (ESA)

1995[19]. In $v = 2006q1$ the production approach was introduced as the principal method for quarterly GDP estimation. In $v = 2007q3$ the estimation method of various annual GDP components changed which greatly affected aggregate GDP. The seasonal adjustment methodology was fundamentally altered in $v = 2009q1$ with the transition from direct to indirect adjustment of GDP. The transition to NOGA[20] 2008 occured in $v = 2012q3$. Furthermore, $v = 2014q3$ will feature the adjustment to the ESA 2010 for all Swiss QNA components.

Minor Changes in the QNA Estimation Methods

In the QNA, the statistical methods and indicators employed for temporal disaggregation are under constant monitoring and may be subject to changes. For example, a discontinuation of any indicator demands a replacement, which may feature slightly different quarterly dynamics or somewhat different seasonal patterns. Hence, the model used for the temporal disaggregation or seasonal adjustment may be changed as well.

Technical Reasons

Even without revisions to the yearly aggregates or the quarterly indicators and without methodological changes of the QNA estimation, slight revisions occur due to the nature of the econometric techniques used for temporal disaggregation. For example, any new data point of a not seasonally adjusted series will lead to new regression coefficients in the model used for seasonal adjustment and, thus, the resulting seasonally adjusted figures will be slightly different for the whole time span of the series.

[19] In the course of 2004 and 2005, the QNA was further revised in the context of the transition to ESA 1995. The transition also included the introduction of the chain-linking method for the calculation of real values.

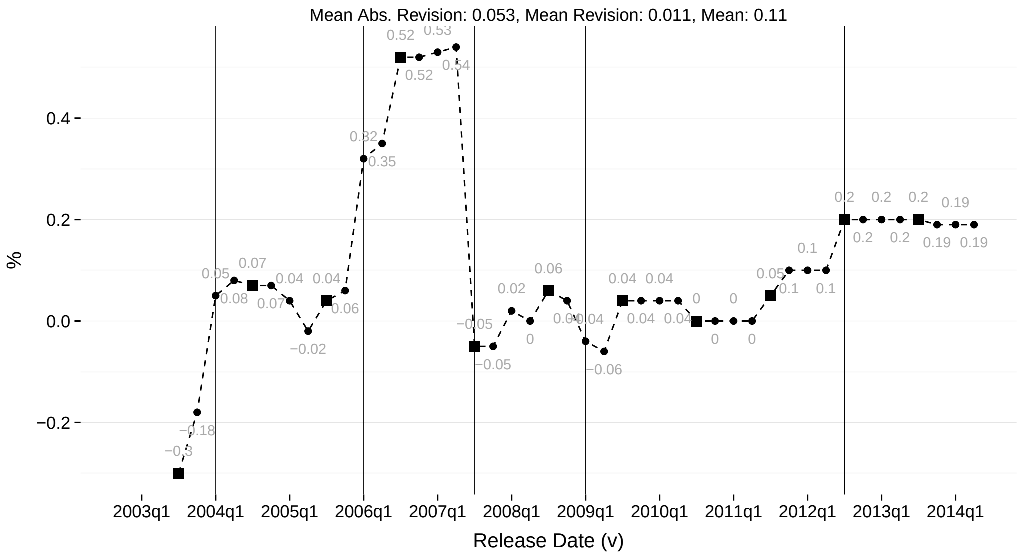
[20] Nomenclature Générale des Activitées Économiques - the Swiss counterpart to the Statistical Classification of Economic Activities in the European Community (NACE)

2.3.2 Revisions to GDP

Revisions to national accounts data can be large and might have great implications on policy making as well as applied economic work. Switzerland is no exception as Figure 2.1 shows. The graph depicts $\Delta gdp_{2003q2}^{v=2003q3-2014q2}$ on the vertical axis and the vintage time index on the horizontal axis. Hence, the dots represent releases of GDP growth rates for the quarter $t = 2003q2$ and vintages $v = 2003q3 - 2014q2$. This corresponds to one row of the triangular matrix in Table 2.1, however, we use growth rates instead of the levels presented in this table. As can be seen, releases of Δgdp_{2003q2} vary within a range of almost 1 percentage point. Even changes in the sign occur several times. In particular, revisions can be strikingly high when a benchmark revision occurs (vertical lines). In contrast, all other revisions are rather small and usually peak when the QNA releases coincided with ANA releases.

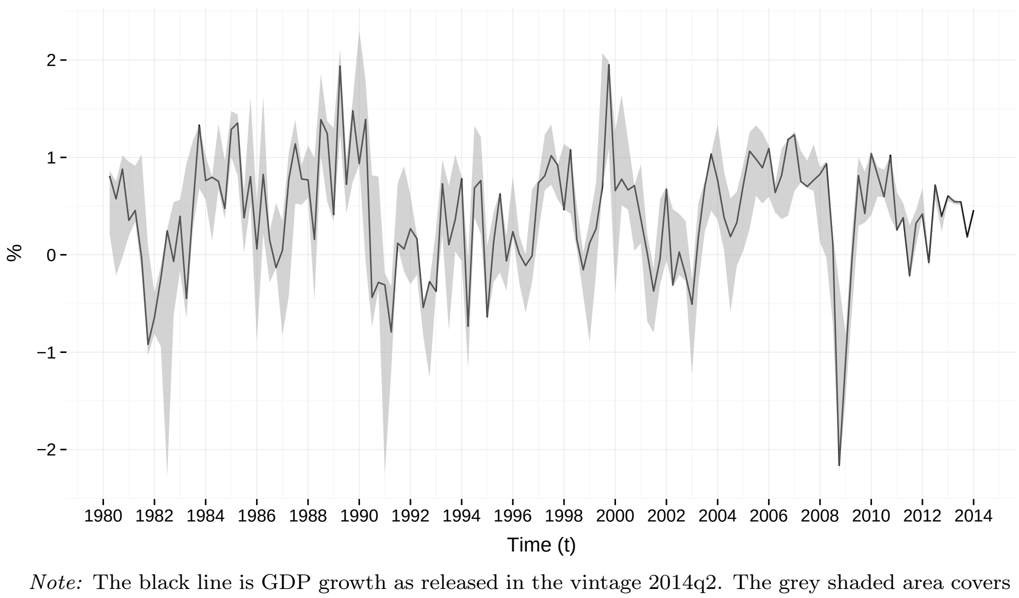
Figure 2.2 shows further that this particular quarter is no exception (this time, the horizontal represents the usual time axis). Whereas the black line represents $\Delta gdp_{1980q1-2014q1}^{2014q2}$, the grey shaded area depicts the range of all published growth rates, $\Delta gdp_t^v \, \forall \, t, v$. Consistent with the conclusions in Siliverstovs (2013), it appears that the course of the business cycle gets not altered fundamentally by the revisions, although the growth rate for a particular quarter may be subject to substantial revisions. Nevertheless, it becomes also clear that revisions can by no means be neglected in Switzerland. For example, both the timing and scope of recessions and expansions may change, resulting in pictures such as Figure 2.1 where a quarterly estimate may indicate a contraction in an early release, and an expansion in a later release, or vice versa. However, revisions should not be interpreted as errors as long as they are not systematically biased. Instead, they indicate the incorporation of newly available or revised and therefore more precise source data, or the adoption of updated concepts or methods.

Figure 2.1: Revision History, 2003q2 (44 Releases)



Note: GDP growth rate of the second quarter 2003 as released in different vintages. Vertical lines correspond to first releases after major methodological revisions. Dots correspond to regular QNA releases. Squares correspond to QNA releases which coincided with ANA releases.

Figure 2.2: Revision History, Swiss GDP



Note: The black line is GDP growth as released in the vintage 2014q2. The grey shaded area covers the range of all growth rates released for a particular time across all vintages. Since for the latest quarter (2014q1) there exists just one vintage (2014q2), the shaded area reduces to a point.

2.3.3 Comparison with other Countries

Table 2.3 shows a comparison of some statistics related to revisions to quarterly GDP in Switzerland, the United States, the Euro Area and Japan. We look at two definitions of revisions. First versus last revisions are given by

$$\begin{aligned} rev_t^{f\ vs\ l} &= \Delta gdp_t^V - \Delta gdp_T^v, \qquad (2.2)\\ V = 2014q2, \quad v &= 2002q4, \dots, 2014q1 \\ t &= 2002q3, \dots, 2013q4 \end{aligned}$$

where V stands for the last vintage available in our data set, T stands for the last observation in the usual time dimension of any given vintage v and Δ takes differences over t, $\Delta x_t^v = x_t^v - x_{t-1}^v$. Therefore, the vector $\Delta gdp_T^{2002q4-2014q1}$ contains all initial (first) releases of a quarterly growth rate and the vector $\Delta gdp_{2002q3-2013q4}^{2014q2}$ contains the most recent estimate of all these quarterly growth rates. This corresponds to comparing the diagonal in Table 2.1 with the last column (in growth rates). First versus second revisions are defined as

$$rev_t^{f\ vs\ s} = \Delta gdp_{T-1}^{v+1} - \Delta gdp_T^v. \qquad (2.3)$$

Hence, the vector of second releases, $\Delta gdp_{T-1}^{2002q4-2014q2}$, is compared with the vector of first releases, $\Delta gdp_T^{2002q4-2014q1}$.

We only consider quarters where initial releases are available in our data set or, put differently, we only use the triangular part of the GDP matrices, $t = 2002q3 - 2014q1$. As a note of caution, Figure 2.1 suggests that large revisions usually occur during benchmark revisions. The number and timing of such revisions may differ across countries and, hence, statistics on data revisions should be compared carefully.

First versus last revisions are lowest in the EA 17. For the US, Japan and Switzerland the extent of the revisions looks similar. In the case of first versus second revisions, this picture does not change fundamentally. However, the first row of Table 2.3 shows that GDP growth rates have been different in the four countries and, therefore, revisions should be compared cautiously.

Table 2.3: Comparison of Revisions to GDP-Growth across Countries.

	CH	US	EA 17	JP
Mean growth of last vintage	0.458	0.4632	0.2054	0.2289
Standard dev. of last vintage	0.6123	0.6666	0.6847	1.1961
First vs. Last				
Mean revision	0.0914	-0.1086	0.0199	-0.0839
	(0.0663)	(0.0497)	(0.028)	(0.0785)
Mean abs. revision	0.3183	0.2806	0.1478	0.3949
Max. abs. revision	1.842	0.8959	0.4099	1.182
First vs. Second				
Mean revision	0.0245	-0.0086	0.0012	-0.0232
	(0.0292)	(0.0163)	(0.0115)	(0.0463)
Mean abs. revision	0.1395	0.0667	0.0543	0.2381
Max. abs. revision	0.6249	0.3681	0.2534	0.9627

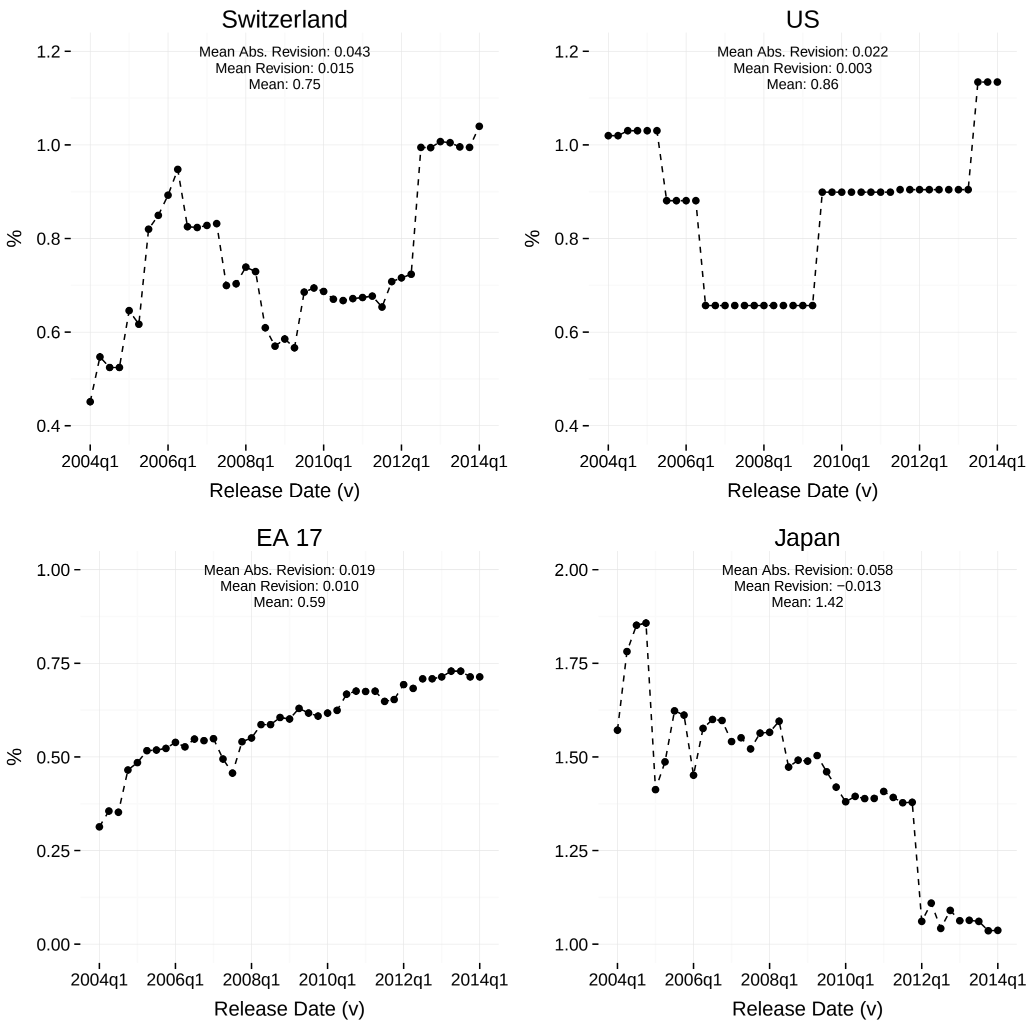
Note: Standard deviation of revisions in parentheses.

These results are broadly speaking in line with Faust, Rogers and Wright (2005) who also find mean absolute revisions to be higher in Japan than in the US but look at an earlier time span. Giannone et al. (2012) focus on a more similar period and find revisions affecting GDP to be lowest in the E.A., somewhat higher in the US and highest in Japan.

Moreover, we find that the dynamics of revisions are different across countries. Figure 2.3 shows GDP releases for 2003q4 for every country. Whereas Δgdp_t got revised continuously in Switzerland, the E.A. and Japan, this is not true for the US There, revisions take place less frequently, but are generally larger when they occur. For example, looking at the upper left panel (Switzerland) and the upper right panel (US), it appears that both mean revisions and mean absolute revisions are higher in Switzerland. However, the range of releases $(max(\Delta gdp_t^v) - min(\Delta gdp_t^v))$ is approximately the same in both countries and, hence, looking at first versus last revisions in Table 2.3 results in similar mean revisions and mean absolute revisions.

This small excercise shows that if one is to evaluate revisions and compare them internationally, it may be important to bear in mind how the figures are computed by the statistical agencies and what particular revision policy they adhere. Revision statistics and hence statistical tests can be influenced by particular features of the data generating processes.

Figure 2.3: Revision History, 2003q4, Cross-Country Comparison.



Note: All panels show the country's GDP growth rate of the fourth quarter 2003 as released in different vintages.The statistics depicted in the upper center of each panel only involve data for 2003q4. I.e., revisions are calculated as $rev^{v}_{2003q4} = \Delta gdp^{v+1}_{2003q4} - \Delta gdp^{v}_{2003q4}$.

2.4 Conclusion

The accuracy of initial estimates of economic variables and the importance of data revisions can only be evaluated by using appropriate and comprehensive real-time data sets as indicated by Eurostat (2013). Moreover, the literature strongly advises the use of real-time data for economic forecasting and policy evaluation if large data revisions occur. Furthermore, it has been shown that it is advisable to check the robustness of empirical economic results across different data vintages if data revisions are present and large. This paper describes such a data set for Switzerland. It contains many important economic variables with quarterly frequency that are subject to revisions. The excercise in the second part of this paper indicates that the magnitude of revisions to quarterly national accounts data of Switzerland are comparable to those of other countries. However, one should be aware of particularities of the revision policies in different countries and their impact on the statistical analysis. Finally, the absence of any revisions would not necessarily be a sign of quality as the inclusion of newly available or revised data increases the precision of the quarterly estimates.

Appendix

2.A Data overview

Table 2.A.1: Variables Contained in the Data Set

Variable	Specification	Sources	Retropolated / Forecasted Vintages
GDP	real, sa	OECD, SECO	$GDP^{2004q1}_{1980q1-1989q4}$, $GDP^{2006q2-2008q1}_{1980q1-1980q4}$
Government Consumption	real, sa	OECD, SECO	$GC^{2004q1}_{1980q1-1989q4}$, $GC^{2006q2-2008q1}_{1980q1-1980q4}$
Private Consumption	real, sa	OECD, SECO	$PC^{2004q1}_{1980q1-1989q4}$, $PC^{2006q2-2008q1}_{1980q1-1980q4}$
Gross Investment	real, sa	OECD, SECO	$INV^{2004q1}_{1980q1-1989q4}$, $INV^{2006q2-2008q1}_{1980q1-1980q4}$
Total Exports	real, sa	OECD, SECO	$EXP^{2004q1}_{1980q1-1989q4}$, $EXP^{2006q2-2008q1}_{1980q1-1980q4}$
Total Imports	real, sa	OECD, SECO	$IMP^{2004q1}_{1980q1-1989q4}$, $IMP^{2006q2-2008q1}_{1980q1-1980q4}$
GDP Deflator	sa	OECD, SECO	$DEFL^{2004q1}_{1980q1-1989q4}$, $DEFL^{2006q2-2008q1}_{1980q1-1980q4}$
Consumer Price Index	nsa	OECD, SFSO	$CPI^{2013q4-2014q2}_{1979q1-1982q4}$
	sa		
Nominal Interest Rate		SNB	$I^{2002q4-2014q2}_{1980q1-1988q4}$
Real Interest Rate			
Employment (fte)	nsa	SECO, SFSO	$EMPL^{2011q2}_{2011q1}$, $EMPL^{2011q3}_{2011q1-2011q2}$
	sa		
Unemployment Rate	nsa	SECO, SFSO	
	sa		
Registered Unemployed	nsa	SECO	
	sa		
Foreign GDP	real, sa	OECD, SECO	
GDP, USA	real, sa	FRED, OECD	
GDP, Eurozone 17	real, sa	Eurostat, OECD	$GDP.EA^{2007q2-2013q3}_{1980q1-1994q4}$
			$GDP.EA^{2008q3}_{2008q2}$, $GDP.EA^{2008q4}_{2008q2-2008q3}$
GDP, Japan	real, sa	Cabinet Office, OECD	$GDP.JP^{2005q1-2009q2}_{1980q1-1993q4}$, $GDP.JP^{2012q1-2014q2}_{1980q1-1993q4}$

2.B Construction of the Unemployment Rate

Deviating from the official calculation in order to mitigate the upward bias, the unemployment rate was calculated as

$$ur_t^v = \frac{Registered\ Unemployed_{\ t}}{Working\ Population_{\ t}^{\ v}} \tag{2.4}$$

where the working population, WP_t^v, is calculated as

$$\underbrace{WP_t^{v<2003q1}}_{\equiv WP^1} = \begin{cases} WP_{1980} + (\frac{WP_{1990}-WP_{1980}}{40})(t-T_0), & \text{if } T_0 \equiv 1980q1 \leq t < 1990q1 \\ WP_{1990}, & \text{if } t \geq 1990q1 \end{cases}$$

$$\underbrace{WP_t^{2003q1\leq v<2012q4}}_{\equiv WP^2} = \begin{cases} WP_t^1, & \text{if } t < 1990q1 \equiv T_1 \\ WP_{1990} + (\frac{WP_{2000}-WP_{1990}}{40})(t-T_1), & \text{if } 1990q1 \leq t < 2000q1 \\ WP_{2000}, & \text{if } t \geq 2000q1 \end{cases}$$

$$WP_t^{2012q4\leq v} = \begin{cases} WP_t^2, & \text{if } t < 2000q1 \equiv T_2 \\ WP_{2000} + (\frac{WP_{2010}-WP_{2000}}{40})(t-T_2), & \text{if } 2000q1 \leq t < 2010 \\ WP_{2010}, & \text{if } t \geq 2010 \end{cases}$$

2.C Construction of the Proxy for Foreign GDP

$$
\begin{aligned}
NGDP_t^{f,v} &= NGDP_t^{ea,v} + NGDP_t^{us,v} + NGDP_t^{jp,v} \\
\Delta gdp_t^{f,v} &= \frac{NGDP_t^{ea,v-1}}{NGDP_t^{f,v-1}} \Delta gdp_t^{ea,v} + \frac{NGDP_t^{us,v-1}}{NGDP_t^{f,v-1}} \Delta gdp_t^{us,v} \\
&+ \frac{NGDP_t^{jp,v-1}}{NGDP_t^{f,v-1}} \Delta gdp_t^{jp,v} \quad \forall t < T \\
\Delta gdp_t^{f,v} &= \frac{NGDP_{t-1}^{ea,v-1}}{NGDP_{t-1}^{f,v-1}} \Delta gdp_t^{ea,v} + \frac{NGDP_{t-1}^{us,v-1}}{NGDP_{t-1}^{f,v-1}} \Delta gdp_t^{us,v} \\
&+ \frac{NGDP_{t-1}^{jp,v-1}}{NGDP_{t-1}^{f,v-1}} \Delta gdp_t^{jp,v} \text{ for } t = T
\end{aligned}
$$

2.D Seasonal Adjustment

Most series get seasonally adjusted by the data producing agencies. Whenever this was not the case, we used the R-interface by Sax (2014) which provides an easy accesses to the new X-13ARIMA-SEATS seasonal adjustment method described in U.S. Census Bureau (2013). We use the automated procedure without outliers. Table 2.D.1 provides details about the models which got applied by the automated procedure. One may raise the objection that X-13ARIMA-SEATS is a very new technology and could not be accessed in real-time. While this is true, it remains impossible to guess what exact seasonal adjustment method might have been applied at a particular point in time in the past. However, the automated procedure for model selection keeps the real-time aspect of seasonal adjustment to a certain extend. Furthermore, the not seasonally adjusted series are also made available in the data set.

Table 2.D.1: Seasonal Adjustment Methods

Variable	Vintage	SARIMA-Model
Consumer Price Index	$2002q4$	$(1,1,0)(1,0,0)$
	$2003q1 - 2003q2$	$(1,1,1)(1,0,0)$
	$2003q3 - 2006q3$	$(1,1,0)(1,0,0)$
	$2006q4$	$(2,1,0)(1,0,0)$
	$2007q1 - 2007q3$	$(1,1,0)(1,0,0)$
	$2007q4 - 2014q2$	$(1,1,0)(0,1,1)$
Employment (fte)	$2002q4 - 2009q1$	$(0,1,0)(0,1,1)$
	$2009q2 - 2012q3$	$(0,1,2)(1,0,0)$
	$2012q4 - 2014q2$	$(0,1,2)(0,1,0)$
Unemployment Rate	$2002q4 - 2003q4$	$(1,1,0)(1,0,0)$
	$2004q1 - 2009q4$	$(1,1,1)(1,0,0)$
	$2010q1$	$(2,1,0)(1,0,0)$
	$2010q2$	$(2,1,2)(1,0,0)$
	$2010q3 - 2010q4$	$(2,1,1)(0,1,0)$
	$2011q1 - 2011q2$	$(2,0,0)(0,1,0)$
	$2011q3$	$(2,1,1)(1,0,0)$
	$2011q4 - 2014q2$	$(2,0,0)(0,1,0)$
Registered Unemployed	$2002q4 - 2003q4$	$(1,1,0)(1,0,0)$
	$2004q1 - 2009q4$	$(1,1,1)(1,0,0)$
	$2010q1$	$(2,1,0)(1,0,0)$
	$2010q2$	$(2,0,1)(0,1,0)$
	$2010q3 - 2011q4$	$(2,1,1)(0,1,0)$
	$2012q1 - 2013q1$	$(2,0,0)(0,1,0)$
	$2011q4 - 2014q2$	$(2,1,1)(0,1,0)$

Chapter 3

Revisions to QNA Data: Are Swiss National Accountants Getting it Right?

Ronald Indergand

Summary: In the 2006 release, annual GDP growth in Switzerland was announced at 1.45% on average for the period 1995-2005. In 2014, average growth for the same period was reported at 1.75%. In this paper, I document the properties of revisions to Swiss National Accounts variables and test for inefficiencies in preliminary data announcements using standard forecast rationality tests. I carefully distinguish between routine revisions and benchmark revisions, where the latter stand for large methodological and definitional changes in the measurement system. Such changes may affect the properties of the time series considerably even in the distant past. In the case of GDP, first releases turn out to be efficient forecasts of all later releases. For some expenditure-side components however, extreme initial releases tend to be revised towards the mean after several quarters. This result corresponds to the findings for other countries and variables. In addition, initial releases of government consumption are positively biased, that is, the preliminary estimates significantly overstated growth rates of later releases during the past decade.

Large parts of this paper were written during a research visit at the Real-Time Data Research Center of the Federal Reserve Bank of Philadelphia. Particular thanks go to Tom Stark for many useful discussions. I also thank Isabel Martinez and Jonathan Wright for helpful comments. I gratefully acknowlege financial support from the Berne University Research Foundation as well as the Irène und Max Gsell Stiftung for this project. I was also supported by the State Secretariat for Economic Affairs, Switzerland (SECO). The views and opinions expressed in this paper are solely mine and not necessarily those of the SECO.

3.1 Introduction

Revisions to macroeconomic data releases are known to be large and they create problems for data users. Policy makers for example are faced with higher uncertainty and thus a more complex decision making process. Moreover, forecast models are sensitive to data revisions (see for example Croushore, 2011*a*). Finally, empirical work in general may deliver different results based on real-time data as opposed to using latest-available vintages which corresponds to the current practice (Croushore and Stark, 2003; Johnson et al., 2013).

Irrespective of their importance for research or policy making, revisions should fulfill certain properties, i.e., they should be well-behaved. This includes a zero mean and being uncorrelated with past information or, put differently, initial data announcements should be rational forecasts of later estimates. In many cases, these conditions are not met. For example, Faust, Rogers and Wright (2005) find predictability of GDP revisions for most G-7 countries. Similarly, Aruoba (2008) finds predictability of revisions for many economic variables in the US and Garratt and Vahey (2006) find the same in the UK. These findings may or may not point to a failure of the statistical agencies.[1] Certainly, however, the implications for policy makers and data producers alike can be very important.

In this article, I examine whether preliminary data announcements in Switzerland are rational forecasts of later releases and if not, I conjecture on the potential sources for the inefficiency.

The first part of the paper documents several characteristics of revisions to Swiss national accounts data. I thereby add to the results in Cuche-Curti, Hall and Zanetti (2008) and Bernhard (2016) for GDP, and extend it to other variables. Revisions in general and benchmark revisions in particular may change properties of the time series considerably even for periods a decade or more in the past. For several GDP-components, it is not unusual that average

[1] A bias may arise from issues not under control of the data producing agency. Moreover, as Aruoba (2008) notes, the agency may create inefficiencies by avoiding other problems, not considered by this literature. Finally, it may take time in some cases to discover existing biases. However, they have to be uncovered first before the agency can adapt.

annual growth rates get revised by up to 0.5 percentage points (or by up to 0.1 percentage points for average quarterly growth) in later data releases. For example, the 2003q1-release of GDP indicated that the Swiss economy experienced an average annual growth of 0.80% during $t = 1990 - 2002$. The Japanese economy grew at 1.26% during this period, greatly exceeding average growth in Switzerland. However, according to the 2015q1 release for the same period, Switzerland grew at an average rate of 1.13% whereas Japan grew by 0.96% per year, reversing the order. This shows that revisions are not only of great importance for the short-term assessment of the business cylce but, in fact, can have implications on how we evaluate medium or long run economic performance across countries. This example may be even more striking in light of the fact that GDP revisions are relatively small compared to other variables considered in this paper.

The second part of the paper employs standard forecast efficiency regressions in order to test for irrationality of first releases (throughout this paper, I use the terms irrational, suboptimal and inefficient synonymously). I distinguish between several definitions of revisions depending on the data production process in Switzerland. As the recent literature suggests, it is crucial that the revisions in question are not contaminated by the effects of benchmark releases (see for example Keane and Runkle, 1990, Swanson and Van Dijk, 2006 or Aruoba, 2008) since these often involve changes of definitions, concepts and methods. So far, studies for Swtzerland did not account for the effect of such large methodological revisions. To do so, I exclude all observations (revisions) from the analysis which involved benchmark releases of Swiss quarterly national accounts (QNA) data.

Reassuringly, I find that initial releases of Swiss GDP are efficient forecasts of all later releases, confirming the results in Cuche-Curti, Hall and Zanetti (2008), Siliverstovs (2011) and Bernhard (2016).[2] For government consumption on the other hand, I find that preliminary data announcements consistently overestimated later growth rates in the past. Furthermore, and contrasting the finding in Cuche-Curti, Hall and Zanetti (2008), neither Bern-

[2] Siliverstovs (2011) classifies GDP releases as news but shows that Business Tendency Surveys may be able to predict future revisions.

hard (2016) nor I find that early releases tend to be smoother than later releases. On the contrary, all expenditure-side components see their initial announcements revised toward the mean after several quarters. That is, high absolute growth rates tend to be revised down in later releases. These findings correspond to the results in Faust, Rogers and Wright (2005) who find mean-reversion for preliminary GDP releases of several other countries. Garratt and Vahey (2006) report the same finding for expenditure-side GDP components as well as many other variables such as retail sales, industrial production or the money stock in the UK. Adding additional variables to the forecast efficiency equations indicates that mean reversion is the most important force behind the predictability of revisions since not much predictive power gets added in most cases.

The remainder of this paper is organized as follows. Section 2 describes the estimation process of national accounts data in Switzerland and presents some stylized facts about the variables in real-time. Section 3 presents some theoretical background on testing data revisions and defines the revisions in question. Section 4 tests for the predictability of these revisions and provides some conjectures about the reasons for inefficient preliminary releases. Section 6 concludes.

3.2 Stylized Facts

3.2.1 National Accounting in Switzerland

National accounts in Switzerland are computed both by the Swiss Federal Statistical Office (SFSO, annual data) and by the State Secretariat for Economic Affairs (SECO, quarterly data). New data for the annual national accounts (ANA) along with revisions of the previous two years are usually released in the third quarter of any year by the SFSO. To compile these statistics the SFSO makes use of representative surveys for final production, intermediate inputs and a broad database for the expenditure side components. New QNA data gets usually released about 2 months after the end of any quarter by the SECO. With each release, the past quarterly data gets revised, in particular the most recent values.[3] Contrary to the annual domain, the available data universe in Switzerland is much smaller on a quarterly frequency. Therefore, the SECO relies on econometric techniques for temporal disaggregation[4] that maintain consistency with the ANA data, that is, quarterly figures add up to the annual counterpart. Therefore, whenever an annual figure gets released or revised, all quarterly values are to be adjusted to those new figures.

Unfortunately, national accounts data is relatively short in Switzerland. At the time of writing, annual values available on the SFSO-homepage started in 1995, compared to 1929 in the case of U.S. data for example, published by the Bureau of Economic Analysis. The SECO retropolates and publishes ANA and QNA data until 1980q1 using quarterly indicators and annual growth rates that were published by the SFSO under an older measurement regime and later discarded when a new measurement system was adopted. Until release 2014q2, annual data back to 1990 and quarterly data back to 1964q1, again retropolated using older data, were available but then discontinued. As a consequence, in the real-time dataset collected by Indergand and Leist (2014), figures further back than 1990 are usually retropolations and should be compared very carefully. I will therefore restrict the subsequent analysis to the time period after 1990.

3.2.2 The Data in Real-Time: Stylized Facts

This section outlines notation and provides some basic facts about growth rates of several variables in real-time, thereby adding to the work of Cuche-Curti, Hall and Zanetti (2008) and Bernhard (2016), and extending it to other variables. As previous work and the example below show, revisions can be so large that every policy maker and applied economist should pay attention to this source of data uncertainty.

Let X_t^v denote the realization of a variable's level at t that got published at release date v by the statistical agency. The natural logarithm is denoted by small letters, x, and Δ takes first differences over t. For any given vintage

[3] For an overview about the causes of such revisions, see Indergand and Leist (2014).
[4] See Sax and Steiner (2013) and the references therein for an overview.

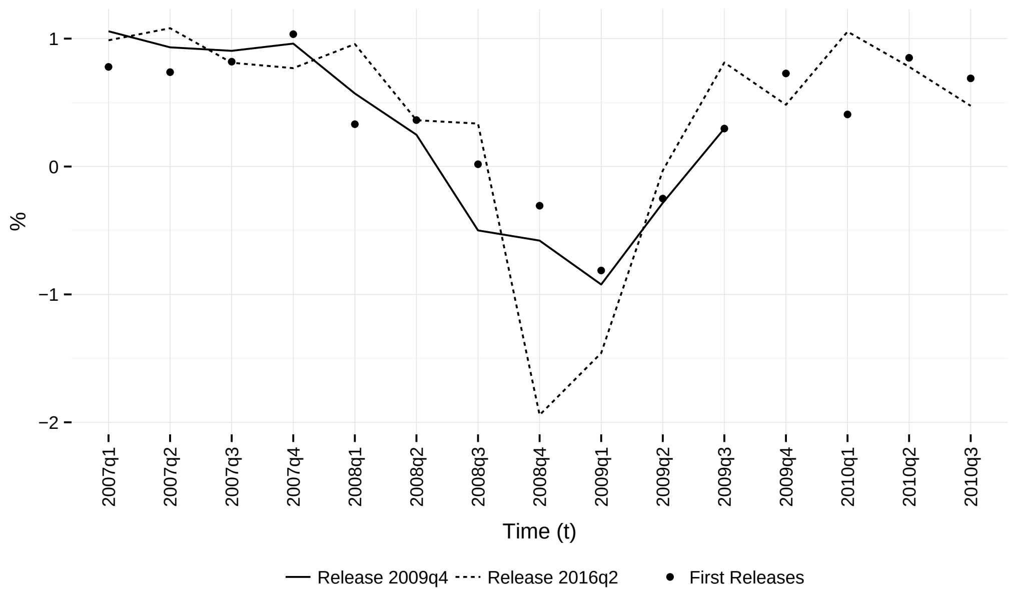
v, T^v stands for the last (most recent) observation available. Collecting $x_{T^v} \forall v$, the vector x_T contains all initial announcements of a variable, i.e., the diagonal elements of the real-time triangular matrix.[5]

As an example, Figure 3.1 shows Swiss GDP growth during the great recession, $t = 2007q1 - 2010q3$, recorded at different points in time. Throughout the paper quarter-on-quarter growth rates are used (not annualized). The dots represent initial releases, Δgdp_T, that is, as GDP was announced by the SECO the first time for each quarter. The solid line shows gdp_t^{2009q4} (the 2009q4 release) and the dashed line gdp_t^{2016q2} which corresponds to the most recent vintage at the time of writing. As can be seen, the begin of the recession and its through shifted in time in opposite directions: According to the most recent release, the recession started later but reached its through earlier. Furthermore, according to the latest estimate, the recession appears to be strikingly more severe at its through as initial releases were suggesting. In fact, the first release for $t = 2008q4$, $\Delta gdp_{2008q4}^{2009q1}$, got revised downward by 1.6 percentage points to $\Delta gdp_{2008q4}^{2016q2} = -1.9\%$, dwarfing the initial announcement of -0.3%.

Figure 3.2 indicates that this is not an unusual example with regard to the size of GDP revisions. The figure depicts box-whisker plots based on all releases, $v = 2002q4 - 2016q2$, for a given quarter during $t = 1990q1 - 2016q1$. The whiskers cover the range whereas the boxes consist of the two middle quartiles of all releases for each quarter. Put differently, a box-whisker plot contains all observations on the horizontal of the real-time triangle for a given quarter. For example, in the case of $t = 1990q1$, the calculations are based on a total of 56 releases, from $v = 2002q4$ through $2016q1$. Since the vintages contained in the real-time data set by Indergand and Leist (2014) start in $v = 2002q4$, the box-whiskers after $t = 2002q3$ lose consecutively one observation. Overall, the revisions for a given quarter can be strikingly large. In 22 out of 104 quarters, the range of the releases for quarterly growth exceeds one percent which compares to an average absolute GDP growth of

[5] Real-Time data is usually stored in matrix form with the release time axis on the horizontal and the observation time axis on the vertical. New vintages are appended on the right of the matrix, typically yielding a See Table 1 on p.334 of Indergand and Leist (2014) for an example.

Figure 3.1: The Great Recession in Real-Time

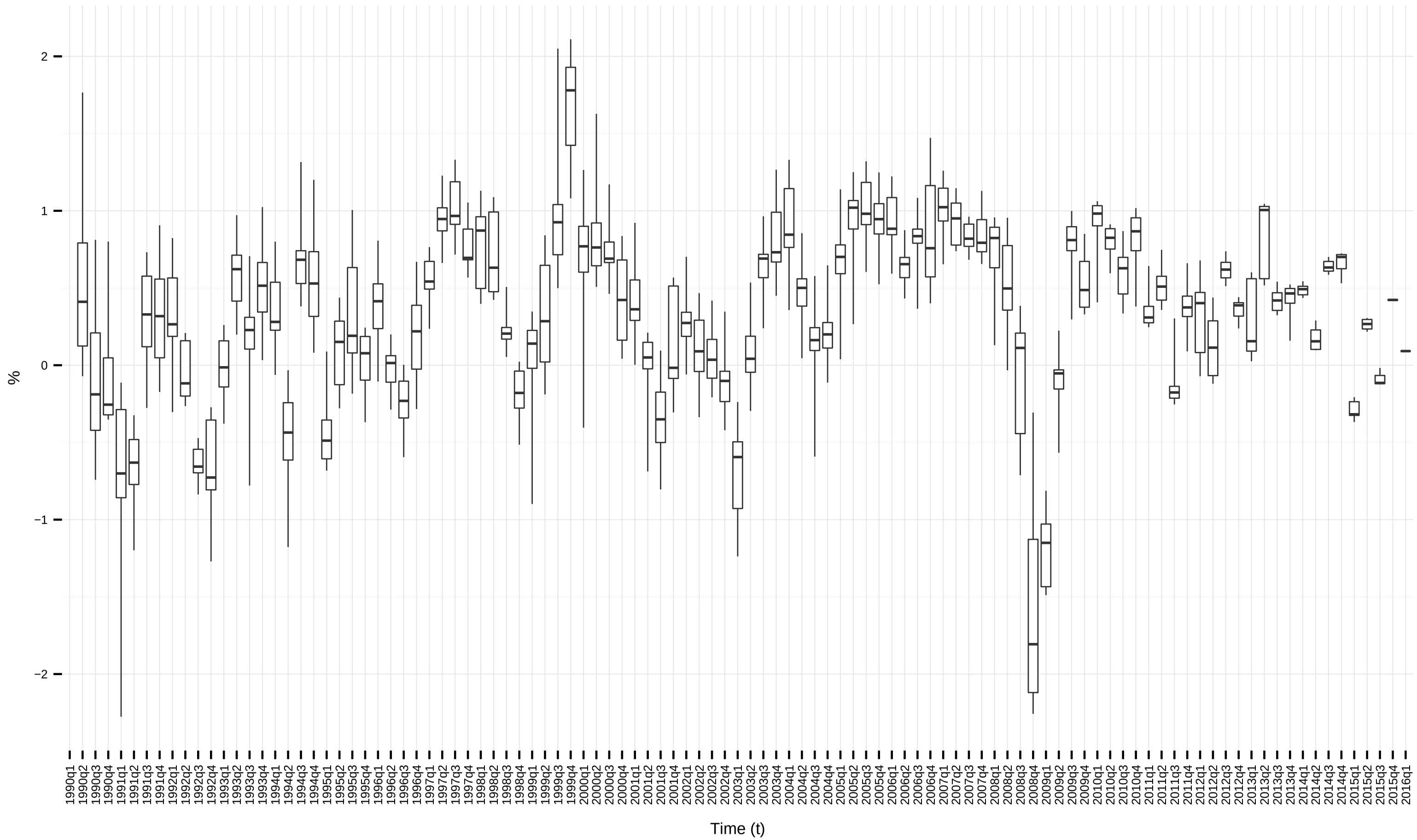


Note: The solid line was released in 2009q4, just after the recession ended. The dashed line was released in 2016q2. The dots represent the first release of any given quarter.

0.6% during the past 25 years (based on the latest available data). Moreover, for more that a third of all observation dates, both negative and positive growth rates had been released. It is also intersting that the box-whiskers in Figure 3.2 become much smaller as one moves to the right of the figure. This indicates that the biggest revisions to quarterly GDP growth in Switzerland occur many quarters after the data has been released for the first time.

Figure 3.3 on the other hand depicts the release date, v, on the horizontal axis. For each vintage, $v = 2002q4 - 2016q2$, the box-whisker plots are based on growth rates during $t = 1990q1 - 2002q3$ (see Appendix A for other variables). The filled boxes correspond to benchmark releases, that is the incorporation of comprehensive methodological or definitional changes. The largest revisions of growth rates usually take place during these fundamental revisions of national accounts. Sometimes the time series also changes substantially with the release of the third quarter of any year, when the QNA release coincides with the ANA release. This is not surprising, since having

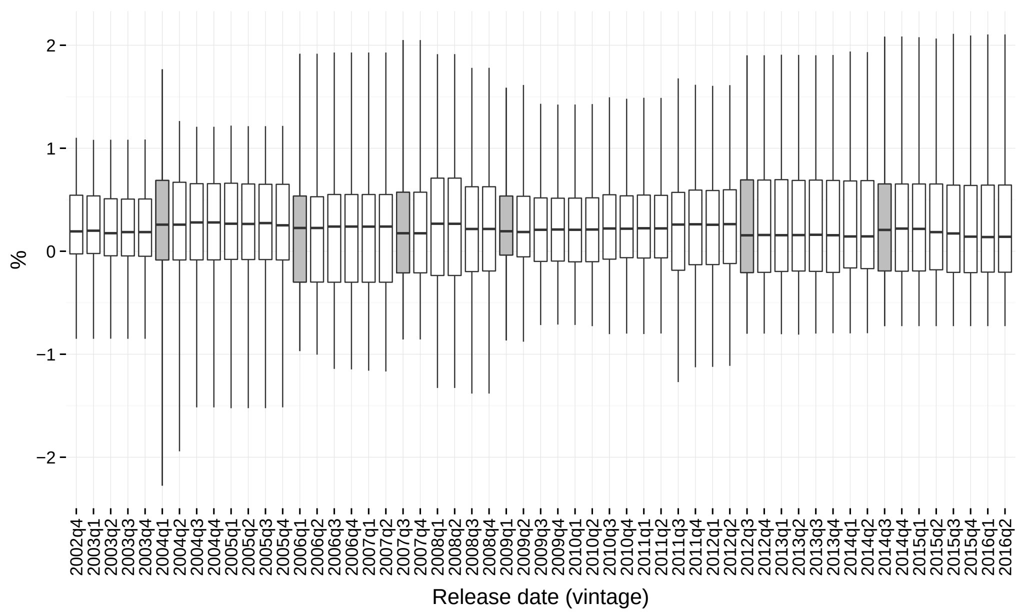
Figure 3.2: GDP growth rates, Box plots for each quarter



Note: For each quarter, $t = 1990q1 - 2016q1$, the figure depicts the median, the range and the two middle quartiles of all releases for the quarterly growth rate. As the real-time data set starts in $v = 2002q4$, all boxplots until $t = 2002q3$ contain 55 observations and consequtively lose one observation for the more recent quarters.

both new and revised annual data at hand, the SECO routinely reviews the indicators and models used for temporal disaggregation. Changes in the disaggregation methods, the indicators used or seasonal adjustment are usually introduced then.

Figure 3.3: Box plots for each vintage: GDP



Note: For each vintage, $v = 2002q4 - 2016q2$, the figure depicts the median, the range and a box containing the two middle quartiles of growth rates for t=1990q1-2002q3. Filled boxes correspond to benchmark releases.

Table 3.1 demonstrates furthermore, that many statistical properties of national accounts time series can be changed profoundly due to benchmark revisions. The table depicts the mean, the standard deviation, the first order autocorrelation coefficient and the number of negative growth rates of GDP and expenditure side components during $t = 1990q1 - 2002q3$ for selected vintages. As the most pronounced changes usually originate from benchmark revisions, the table shows only the figures for those six release dates and the first vintage in the data set. Both the mean and the standard deviation may change substantially for some variables even when the period in question lies more than a decade in the past. For example, mean growth for total investment during $1990q1$ and $2002q3$ was revised from -0.13% (released in $v = 2002q4$) up to 0.21% (released in $v = 2014q3$). Moreover, the autocorrelation structure in all variables considered was often changed during benchmark revisions. In particular, all series showed large, significant autocorrelation until the $2002q4$ release. This pattern got completely altered with the benchmark revision of $2004q1$ when the new European System of Accounts (ESA) 1995 was introduced. After this adjustment, all variables displayed little or no autocorrelation.

It becomes also apparent that revisions to almost all expenditure side components can be much larger than for aggregate GDP. However, this is not necessarily surprising. Before $v = 2006q1$, quarterly GDP was directly estimated, that is, independently of the expenditure-side components (SECO, 2006). Revisions to the components therefore did not automatically translate to the aggregate. In $v = 2006q1$ the production apprach was introduced as a new standard for GDP estimation. Since then, quarterly GDP is measured as the sum of gross value added by producers (industries). The difference between GDP and the sum of the expenditure side components is captured in a statistical discrepancy component which also includes investment in inventories. Note, however, that in the estimation process of the ANA, GDP may be more directly linked to the expenditure side components which of course will also translate to the QNA data to some extend.

Finally, the bottom panel in Table 3.1 shows that the number of negative growth rates also changes substantially across vintages for many variables. This suggests that policy makers may face considerable challenges in assessing whether the economy is in a downturn or not.

Table 3.1 assesses this question somewhat more systematic. The shaded areas mark contractionary periods during $t = 1990q1 - 2003q3$ for the same selection of vintages as in Table 3.1. Panel A applies the unofficial but in Switzerland widely used definition of a recession as a fall of GDP during two successive quarters or more. In Panel B economic downturns are defined according to Harding and Pagan (2002) who implement the algorithm of Boschan and Bry (1971) on the quarterly level. This algorithm uses more sophisticated ad-hoc rules to determine peaks and troughs of the business cycle. They are determined as local minimas and maximas of a univariate time series using constraints with regard to the duration and amplitude of

Table 3.1: Summary statistics for selected vintages

	2002q4	2004q1	2006q1	2007q3	2009q1	2012q3	2014q3
				Mean			
GDP	0.23	0.29	0.28	0.28	0.27	0.27	0.30
GC	0.42	0.37	0.40	0.34	0.32	0.27	0.32
PC	0.28	0.32	0.30	0.32	0.31	0.32	0.30
INV	-0.13	0.03	0.18	0.14	0.14	0.14	0.21
EXP	0.74	0.84	0.85	0.90	0.89	0.91	0.89
IMP	0.72	0.72	0.75	0.82	0.82	0.81	0.84
				Standard deviation			
GDP	0.44	0.69	0.67	0.69	0.53	0.56	0.60
GC	0.45	0.82	0.67	0.79	0.73	1.02	1.13
PC	0.32	0.60	0.59	0.56	0.41	0.44	0.38
INV	1.37	2.53	1.96	2.35	2.35	2.01	1.91
EXP	1.42	2.19	2.05	2.06	2.12	2.05	1.90
IMP	1.85	2.44	2.54	2.68	2.64	2.32	1.82
				First order autocorrelation			
GDP	0.70	0.27	0.18	0.19	0.45	0.27	0.40
GC	0.79	-0.02	0.35	0.22	0.23	-0.26	-0.21
PC	0.69	-0.30	-0.26	-0.22	0.32	0.05	0.03
INV	0.82	-0.13	0.03	-0.14	-0.15	0.06	0.14
EXP	0.66	0.10	0.15	0.13	0.13	0.16	0.13
IMP	0.72	0.07	0.02	-0.03	-0.02	0.06	0.03
				Number of negative growth rates			
GDP	15	14	17	19	14	16	16
GC	8	13	16	16	15	20	19
PC	7	14	13	15	12	11	12
INV	25	27	24	31	31	25	21
EXP	15	15	18	16	16	14	17
IMP	15	20	20	20	20	20	19

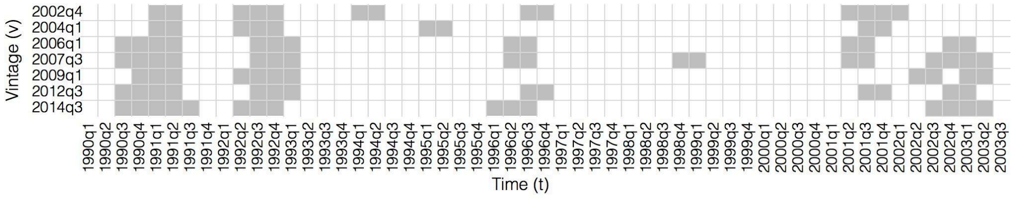
Note: All calculations based on the period t=1990q1-2002q3.

an expansionary or contractionary phase as well as the duration of the full cycle.

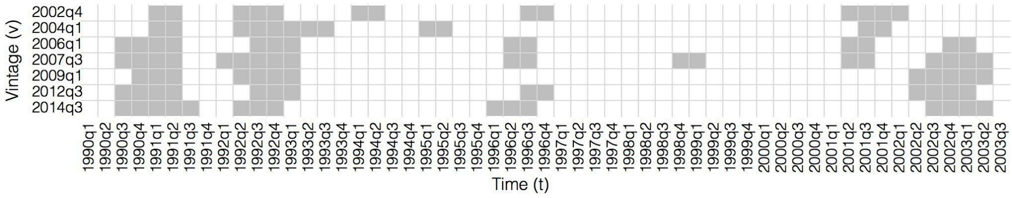
Using both definitions, Table 3.1 indeed confirms that recessions may shift in time, lengthen, shorten, newly appear or disappear altogether as one looks at different vintages of GDP. It is notable that the naive and the more complicated algorithm identify very similar contractionary periods and the effect of data revisions is comparable. For example, looking at vintage $v = 2002q4$, in both panels one finds five recessions in the data whereas only three are to be found if one looks at the $2014q3$ release (or the most

Table 3.2: Recessions in Real-Time

Panel A: At least two negative, consequtive quarters



Panel B: Harding-Pagan algorithm (quarterly version of Bry-Boschan)



Note: Shaded areas indicate recessions measured by Swiss GDP. Panel A defines an economic contraction as two or more successive, negative quarters. Panel B uses the algorithm of Harding and Pagan (2002) to isolate peaks and throughs. The rows represent the vintages which Indergand and Leist (2014) define as benchmark revisions as well as the first vintage in their data set. Note that for $v = 2002q4$, the observations end in $t = 2002q3$.

recent GDP release at the time of writing). The length of the recession at the beginning of the 1990s varies between two and five quarters. During the 1990s and later, several contractionary periods may be detected - or not detected depending on the data vintage at hand.[6]

Reassuringly, most of the shaded areas in Table 3.2 represent periods of relatively weak economic activity in all releases. Nevertheless, this example shows that at least when using these mechanic definitions for recessions, conclusions may be heavily influenced by measurement issues. This suggests that a definition of economic downturns by a convention that takes into account more information than a univariate algorithm might be useful also for Switzerland. This is for example done for the US by the National Bureau of Economic Research and for the Euro Area by the Centre for Economic Policy Research.[7]

[6] Siliverstovs (2013) looks at this question using two alternative business cycle dating techniques. See also Bernhard (2016).

[7] See `http://www.nber.org/cycles.html` and `http://cepr.org/content/euro-area-business-cycle-dating-committee`.

All this suggests that revisions in general and benchmark revisions in particular can alter both the economic interpretation of a time series and its statistical properties profoundly. Preliminary results show that data revisions seem to be at least as important in the Swiss case compared to bigger economies (Table 3.B.1 in the Appendix reproduces Table 3.1 for GDP of the US, the Euro Area and Japan). Hence, applied economists should pay attention to the revision policy of statistical agencies and the effect of data revisions should be taken into account for economic modelling.

Policy makers generally rely on the assumption that preliminary releases of economic variables are as accurate as possible. Given that measurement issues can have such large effects on the properties of economic data, it is very important to know whether the preliminary data releases are efficient, that is, unpredictable. The following sections provide a formal analysis to answer this question for national accounts data in Switzerland.

3.3 Theoretical Background

3.3.1 A Framework for Analyzing Data Revisions

Data revisions can be studied using the following framework. The exposition follows losely Mankiw, Runkle and Shapiro (1984) who were among the first to propose such a framework but many studies have taken a similar approach. For an overview of this literature, see Croushore (2011*a*) or Jacobs and Van Norden (2011).

Let x_t^f be the final value of an economic variable at period t announced at release date f and let x_t^p be its preliminary estimate announced at release date p. The final data equals its preliminary estimate plus an error component:

$$x_t^f = x_t^p + \epsilon_t. \tag{3.1}$$

Two polar cases can be considered with regard to the error term. Under the *news* hypothesis, x_t^p is an efficient (rational, optimal) forecast of x_t^f and ϵ_t reflects new information that appeared after the preliminary estimate was announced. The preliminary data is optimal in the sense that only new information, and no information known at p, can bring it closer to the final value. In this case, the error term (revision) will be orthogonal to the preliminary release, $corr(x_t^p, \epsilon_t) \neq 0$.

Under the *noise* hypothesis, x_t^p is an inefficient (irrational, suboptimal) forecast of x_t^f. In this case, the preliminary release is contaminated by measurement error that is uninformative about the final value. This error gets reduced over time such that x_t^p converges to x_t^f. The error term will be uncorrelated with the final release, $corr(x_t^f, \epsilon_t) = 0$, but correlated with the preliminary release: $corr(x_t^p, \epsilon_t) \neq 0$. The optimal forecast, x_t^*, would then be a transformation of x^p using the information about $corr(x_t^p, \epsilon_t)$.

Hence under the noise hypothesis, running the regression

$$x_t^f = \alpha + \gamma x_t^p + \delta_t \tag{3.2}$$

will result in $\hat{\gamma} \neq 1$ violating forecast rationality. Equation (3.2) is also known as the Mincer-Zarnowitz regression (Mincer and Zarnowitz, 1969). Equivalently, substracting x_t^p from both sides of (3.2) and thus regressing the revision on the preliminary release,

$$rev_t^f = \alpha + \beta x_t^p + \nu_t, \tag{3.3}$$

where $\beta = \gamma - 1$, will result in $\hat{\beta} \neq 0$. In this case, the preliminary data release itself is useful for predicting its subsequent revision and hence is an inefficient forecast of x_t^f.

Of course, x_t^p is also suboptimal if anything else contained in the information set known at the time of the preliminary release, $\mathcal{I}^p$, is useful to predict the final release. Suppose that the statistical agency announces a preliminary estimate that differs from the optimal forecast, x_t^*, such that $x_t^* = bx_t^p + IB$ where I is a matrix of additional variables and B their nonzero weights. Then in the regression

$$x_t^f = \alpha + \gamma x_t^p + I\Gamma + \epsilon_t, \tag{3.4}$$

we will find $\hat{\alpha} = 0$, $\hat{\Gamma} \neq 0$ and potentially also $\hat{\gamma} \neq 1$. Equivalently, using the revision as a dependent variable,

$$rev_t^f = \alpha + \beta x_t^p + I\Psi + \epsilon_t, \tag{3.5}$$

we will find $\hat{\alpha} = 0$, $\hat{\Psi} \neq 0$ and potentially $\hat{\beta} = \gamma - 1 \neq 0$. Finally, note that forecast rationality is also violated if $\hat{\alpha} \neq 0$ in Equations (3.2) - (3.5) or if the unconditional mean of the revisions differs from zero. In the latter case, I label preliminary estimates as biased and in the former cases as inefficient.

Very often, the actual true data of an economic variable will never be known and the revision process may continue for a long time or even indefinitely. Assuming that all revisions move the estimate closer to the true value, the final (latest) release will be the closest proxy of the true data. However, note that also intermediate releases can be tested using the above framework. The following section will thus define the revisions in question and argue why these revisions are an interesting object to study.

3.3.2 Definition of Revisions

Generally, revisions are defined as

$$rev_t^h = \Delta x_t^{v+h} - \Delta x_t^v, \tag{3.6}$$

with $h > 0$. That is, rev_t^h captures the cumulative revisions of a growth rate between v and its later release at $v + h$. Since the revision process may continue indefinitely, there is no clear guidance about which vintage, $v + h$, should be chosen as a reference to evaluate earlier releases. Its choice will both depend on the research question and on the data construction process described above.

Earlier analyses, often due to the lack of suitable data, tended to analyze short-term revisions (low h, for example Mankiw and Shapiro, 1986) and in particular long-term revisions based on the last vintage available (maximal h, see for example the studies listed in Swanson and Van Dijk, 2006).

The latter assumes that every revision brings an observation closer to the true value which it is designed to measure. However, this falls short of acknowledging the fact that measurement concepts can change over time and that certain variables therefore are not measuring exactly the same thing in an early vintage compared to a later release - even if they are meant to measure the same underlying economic variable, such as total investment or total economic activity. More recent contributions therefore emphasize the role of so-called benchmark releases. A benchmark release incorporates a comprehensive revision of the data, that may involve conceptual and definitional changes as well as new estimation methods. It seems reasonable to assume that any data producing agency strives "to make the preliminary data the best possible estimate [...] *under the current measurement system*" (Faust, Rogers and Wright, 2005, p. 406). For example, up to release 2014q2, investment in equipment according to the Swiss national accounts did not include investment in research and development. With the benchmark revision of 2014q3, this and other changes were introduced (SECO, 2014*b*) and as a consequence average nominal growth for investment in equipment during $t = 1995q1 - 2014q1$ was revised from 0.43% to 0.59%. However, even if such a revision of growth rates was perfectly foreseeable to the data producers, they will not attempt to revise the published figures - nor should they - as long as they release the data under the old measurement system. Hence an improper treatment of benchmark revisions could impede conclusions from efficiency tests of first releases. In fact, ignorance of benchmark revisions may strongly affect the results as Keane and Runkle (1990) show in the case of testing for unbiasedness of individual price forecasts and Swanson and Van Dijk (2006) in the case of testing for unbiasedness of initial price estimates. Therefore, Aruoba (2008) suggests that h in Equation (3.6) should be as large as possible but not include benchmark releases (see also Croushore, 2011*a*). Swanson and Van Dijk (2006) on the other hand try to back out benchmark revisions by pre-treating the data. Finally, Garratt and Vahey (2006) deal with this problem by allowing for several structural breaks in the regression equations.

In order to test for inefficiency of the initial releases, I follow Aruoba (2008) and exclude an observation from the analysis whenever a benchmark revision occurs between v and h in Equation (3.6). In order to classify a vintage as a benchmark release, I follow Indergand and Leist (2014) who provide an overview about all comprehensive revisions for both the Swiss ANA and QNA during the covered period.

For all efficiency tests, I focus on four different definitions of revisions due to several reasons. First, this provides robustness checks for the findings. Second, policy makers and applied economists may have interest in different revision types depending on the topic, the research question or the forecast horizon. Third, if forecast efficiency is violated, this will provide information on where to look for the sources of the problem. The four definitions are as follows. Initial revisions,

$$rev_t^{init} = \Delta x_{T^v}^{v+1} - \Delta x_{T^v}^{v}, \tag{3.7}$$

measure the change between the first release of an observation, $\Delta x_{T^v}^{v}$, and its second release, $\Delta x_{T^v}^{v+1}$. Arguably, information about initial revisions may be of particularly high relevance for policy makers since policy decisions are often based on the preliminary data release. If initial revisions were biased, or its standard deviation extraordinarily high, policy makers would be well advised to rely on a transformation of the first release, or in the extreme case even to wait and base their decisions on the second release.

For the second revision type, h corresponds to the last vintage before the extrapolated QNA figures get adjusted to a new ANA release:

$$rev_t^{ana-1} = \Delta x_{T^v}^{ana-1} - \Delta x_{T^v}^{v}. \tag{3.8}$$

ana corresponds to vintages that feature new as well as revised annual data, typically corresponding to the third quarter. Neither rev_t^{init} nor rev_t^{ana-1} are directly contaminated by any effects related to the ANA data.[8]

[8] Note, however, that the ANA figures of course have implications on the models and indicators used for quarterly extrapolation.

Hence, if any bias was detected for these two revision types, one may conclude that first releases of the QNA are inherently inefficient.

The following definition can be used to test whether the preliminary releases are an efficient estimate of the figures that incorporate new annual information:

$$rev_t^{ana} = \Delta x_{T^v}^{ana} - \Delta x_{T^v}^{v}. \tag{3.9}$$

In every *ana* vintage, four extrapolated quarterly observations got adjusted to a new annual observation such that their sum equals the ANA counterpart. Furthermore previous ANA data gets typically revised which also translates into new interpolated quarterly values. Note that a bias in rev_t^{ana} could arise both due to inefficiency of the first QNA or the first ANA releases. The reason is that the most recent annual data has a direct impact on its quarterly extrapolation as well as an indirect impact through the indicators and models, that were used for temporal disaggregation and chosen on the basis of the annual figures. Hence, a bias in rev_t^{ana} could also appear if first ANA releases are an inefficient forecast of later ANA releases.

Finally, ignoring all benchmark releases, I also report results for the case where $v + h$ corresponds to the latest available vintage, V, corresponding to $v = 2016q2$ at the time of writing (final revisions):

$$rev_t^{final} = \Delta x_{T^v}^{V} - \Delta x_{T^v}^{v}. \tag{3.10}$$

This choice bears less relevance for assessing the efficiency of the statistical agencies who produce the data according to a certain measurement convention. Nevertheless, analyzing final revisions provides information on whether these measurement concepts and approaches overall succeed in measuring an underlying variable (such as economic activity, measured by GDP). Of course, this assumes that the underlying variable that is supposed to be measured remains the same, eventough measurement concepts and definitions change. These results may be of interest for decision makers and scholars alike since ideally, they would like to base their decisions or analyses on data that efficiently measures the true state of the economy.

I further check for the robustness of the results using additional definitions. For example, one can test for a bias in first releases compared to the latest vintage under the respective measurement system. In this case $v + h$ includes as many vintages as possible but stops before a benchmark release occurs. Using this definition, however, it becomes impossible to infer wether any bias in first releases stem mainly from the QNA or the ANA estimation process.

Finally, note that the conclusions in this paper do not necessarily bear validity for the future. National accounting methods are continuously evolving. Hence it need not be the case that past predictability of revisions also implies future predictability. Furthermore, as Aruoba (2008) notes, statistical agencies might be avoiding other problems at the expense of the problems outlined by this literature.

In the remainder of this paper, I will test these revisions for all three forms of predictability defined in Section 3.31. First, whether the unconditional mean of the revisions is nonzero. Second, whether the Mincer-Zarnowitz test is violated (Equation 3.3) and third, whether additional information may be used to forecast data revisions (Equation 3.5).

3.4 Are Revisions Predictable?

3.4.1 Unconditional Summary Statistics

Table 3.1 provides summary statistics and tests for unbiasedness of the revisions to GDP, government and private consumption, investment, imports and exports. For each variable the table shows the mean for all four types of revisions defined in Section 3.32. Standard errors are for the hypothesis that the mean is equal to zero are autocorrelation and heteroskedasticity robust (Newey and West, 1987), based on a lag truncation parameter of 4. The table also shows the standard deviation of the revisions, the mean absolute, a standardized mean absolute, the maximum and the minimum revision. For the standardized mean absolute revision, the revisions are divided by the mean absolute of all first releases in order to enhance the comparability across variables.

Table 3.1: Summary statistics of revisions

	rev^{init}	rev^{ana-1}	rev^{ana}	rev^{final}
		GDP		
Mean	0.03	0.03	0.01	0.09
Standard deviation	0.18	0.16	0.25	0.46
Mean absolute	0.12	0.12	0.18	0.35
Standardized mean absolute	0.28	0.26	0.37	0.78
Maximum	0.62	0.38	0.62	0.97
Minimum	−0.53	−0.53	−0.61	−1.64
		Governement Consumption		
Mean	−0.12**	−0.15***	−0.29***	−0.16*
Standard deviation	0.48	0.25	0.58	0.82
Mean absolute	0.29	0.20	0.48	0.67
Standardized mean absolute	0.42	0.27	0.71	0.94
Maximum	1.17	0.23	1.17	1.26
Minimum	−1.83	−0.99	−1.51	−1.97
		Private Consumption		
Mean	0.01	0.02	−0.01	−0.02
Standard deviation	0.16	0.13	0.25	0.28
Mean absolute	0.11	0.10	0.20	0.21
Standardized mean absolute	0.28	0.26	0.51	0.54
Maximum	0.61	0.23	0.61	0.67
Minimum	−0.32	−0.26	−0.46	−0.47
		Total Investment		
Mean	0.03	0.15**	0.15	0.26
Standard deviation	0.62	0.61	0.90	1.60
Mean absolute	0.40	0.43	0.71	1.18
Standardized mean absolute	0.32	0.28	0.58	0.96
Maximum	1.75	1.75	2.61	5.08
Minimum	−1.72	−1.2	−1.41	−5.17
		Total Exports		
Mean	−0.03	0.19	0.14	0.23
Standard deviation	0.80	0.77	1.30	2.12
Mean absolute	0.55	0.53	1.02	1.60
Standardized mean absolute	0.29	0.30	0.50	0.88
Maximum	1.53	2.18	2.71	5.05
Minimum	−2.23	−1.95	−2.23	−5.6
		Total Imports		
Mean	0.14	0.29***	0.10	0.39***
Standard deviation	0.76	0.68	0.97	1.82
Mean absolute	0.53	0.49	0.75	1.47
Standardized mean absolute	0.29	0.24	0.43	0.82
Maximum	2.43	2.43	1.86	4.47
Minimum	−2.04	−0.6	−2.04	−4.42
N	48	35	31	54

Note: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Newey-West standard errors with four lags. Calculations based on the full sample, $t = 2002q3 - 2016q1$, and $v = 2002q4 - 2016q2$

With the exception of government consumption, revisions have been mostly positive on average (though not statistically significant, see below), particularly for final revisions when benchmark releases are not excluded from the analysis. This is consistent with Table 3.1 where an earlier time period, instead of first releases, was considered. Furthermore, this seems also to be the case in other countries. For the USA, Aruoba (2008) does not find a single variable among the many he analyzes for which mean revisions have been negative during 1965 and 2005. For the UK, Garratt and Vahey (2006) find similar results. Likewise, short-term and long-term revisions to GDP have been generally positive in the case of seven big countries prior to 1997 (Faust, Rogers and Wright, 2005).

The standard deviation of revisions and the mean absolute revision increase sharply for all variables in the case of rev^{ana} as well as for rev^{final}. Eventough preliminary releases are associated with more uncertainty than second or third releases, this may not be particularly surprising considering the fact that important new information arrives with every *ana*-vintage. When the relative mean absolute revision is standardized by the mean absolute growth rates of the preliminary releases, it becomes clear that revisions are comparatively less important for GDP, whereas they are most sizeable for government consumption and investment. In case of the latter for example, a standardized mean absolute revision of 0.96 means that the revisions are on average about as big as the mean absolute change of the preliminary growth rates, implying a lot of uncertainty in case of preliminary estimates of total investment.

Turning to the statistical tests of the revisional mean, the most notable result is that the revisions to preliminary releases of government consumption are biased in all cases. For initial revisions, this result may be driven by a few outliers (see the upper left graph in Figure 3.2). However, the result is very clear in the case of the third type of revisions, rev^{ana}, after the quarterly extrapolations got adjusted to new annual data (see the lower left graph in Figure 3.2): Revisions to almost all preliminary releases are either close to zero or negative and the preliminary releases overestimate the revised growth rates by almost 0.3 percentage points. This is particularly striking considering the fact that the overall mean of government consumption based on the latest vintage was little more than 0.3% during this time period.

Moreover and surprisingly, this pattern diminishes (but remains statistically significant on the 5% level) when final revisions, rev^{final}, are considered: The lower right graph in Figure 3.2 suggests that these revisions are much more centered around zero. In fact, excluding the most extreme negative revisions drives the mean of rev^{final} down to zero. This means that the preliminary quarterly extrapolations overestimate growth rates of government consumption in general and clearly are a biased predictor of the intermediate releases that got adjusted to the annual data. However, these intermediate releases get revised back in the direction of the first announcements in later releases.

For investment and import data, the revisions are on average positive and in a few cases statistically different from zero as well. But again, these results are mainly driven by one or two observations (see Figure 3.D.2 and 3.D.4 in the Appendix). For GDP, private consumption and exports on the other hand, all revisions have a mean that is not statistically different from zero.

3.4.2 Baseline Efficiency Regressions

This section turns to the forecast efficiency test given in Equation (3.3). Table 3.2 shows the results for the estimated coefficients together with the adjusted R^2 statistic. It also shows the Wald test statistic for the joint test $\alpha = \beta = 0$. Figure 3.1 and Figure 3.2 provide graphical evidence for GDP as well as government consumption and Figures 3.D.1 - 3.D.4 (Appendix) for the other variables. First releases of growth rates are depicted on the horizontal and revisions to these releases on the vertical, seperately in each panel for one of the revision types defined in Equations (3.7) to (3.10).

In the case of GDP, the first releases and their revisions are uncorrelated in all cases, indicating no inefficiency. This contrasts with results that have been found for GDP of several other countries.[9] However, the picture changes

[9] See Faust, Rogers and Wright (2005), Garratt and Vahey (2006). Note, however, that final revisions to Swiss GDP, rev^{final}, become somewhat forecastable if one excludes

to some extend if one looks at the expenditure-side components of GDP: A generally negative relationship between revisions and preliminary releases appears. The coefficients are small and mostly insignificant in the shorter term (rev^{init} and rev^{ana-1}). But they become much bigger and in almost all cases highly significant after the data has been adjusted to new ANA data (rev^{ana}) and even more so in the very long term (rev^{final}). This means that extreme initial growth rates generally get revised towards their mean, i.e., high initial estimates tend to be revised downwards and low estimates tend to be revised upwards.

Preliminary releases of GDP and its components are forecasts of their later releases and for their estimation, the statistical agencies rely often on econometric inter- and extrapolation methods. Considering the fact that economic forecasts based on regression analysis tend to be smoother than the actual data, it may be surprising that the preliminary releases exhibit excessive variability. However, it is consistent with the noise interpretation of data revisions where the preliminary estimates are polluted with uninformative noise that gets reduced lateron. This result is also very much in line with the findings of Garratt and Vahey (2006) who focus on the UK and find negative and significant coefficients for the same expenditure-side variables as considered in this paper. It is interesting, however, that this pattern is very small at the beginning and becomes much more amplified after several quarters at least in the Swiss context.[10] Section 3.44 outlines possible causes for this result.

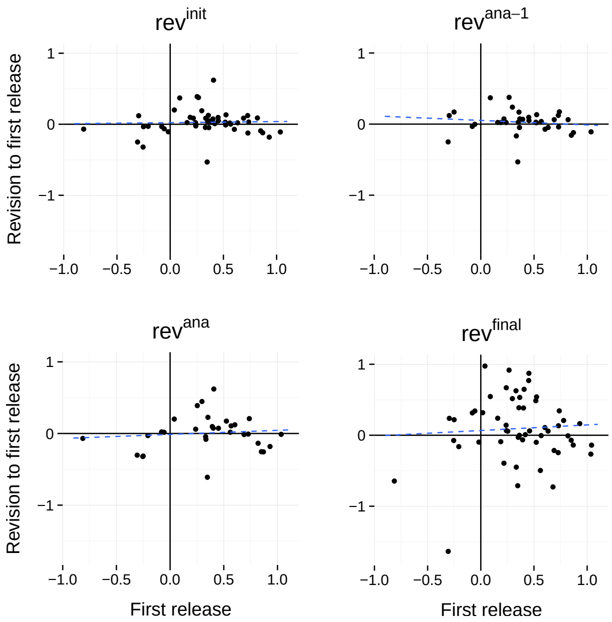
3.4.3 Augmented Efficiency Regressions

If first data announcements are rational forecasts of later releases, no information known at the release date v should be useful to predict the revisions,

most negative preliminary release and the most negative revision from the analysis (the two extreme observations in the lower right panel of Figure 3.1). In this case a sizeable negative coefficient of -0.28, significant on the 5% level, is estimated. The explanatory power remains close to zero nevertheless.

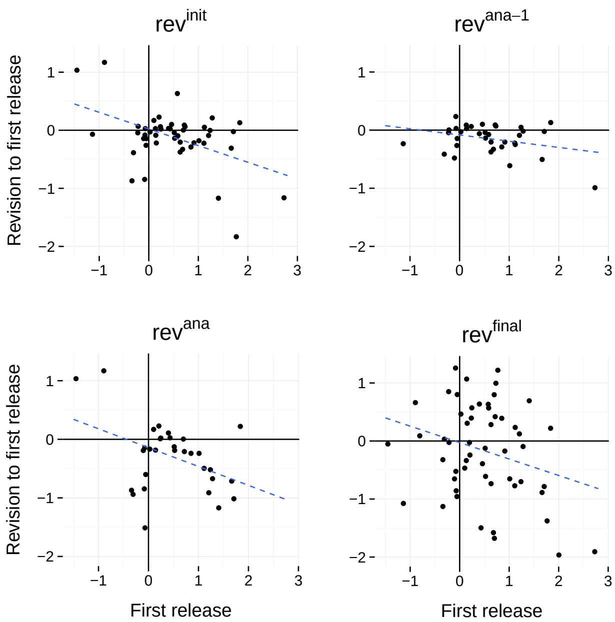
[10] Faust, Rogers and Wright (2005) also find a somewhat weaker (but still strong) predictability when they focus on short-term revisions in the case of GDP. They define short-term revisions as the comparison between the first release and the release 8 quarters later.

Figure 3.1: First Estimates and Revisions, Gross Domestic Product



Note: The dotted line corresponds to an OLS regression of revisions on first releases and a constant.

Figure 3.2: First Estimates and Revisions, Government Consumption



Note: The dotted line corresponds to an OLS regression of revisions on first releases and a constant.

Table 3.2: Baseline Efficiency Regressions (OLS)

	rev^{init}	rev^{ana-1}	rev^{ana}	rev^{final}
		GDP		
Constant	0.02	0.05	−0.01	0.07
First release	0.02	−0.06	0.06	0.08
Adj.R^2	−0.02	−0.01	−0.02	−0.02
W	1.56	1.91	0.84	4.22
		Governement Consumption		
Constant	0.02	−0.08***	−0.14	−0.03
First release	−0.29***	−0.11	−0.32**	−0.28*
Adj.R^2	0.20	0.08	0.15	0.06
W	21.30***	30.96***	23.61***	6.29**
		Private Consumption		
Constant	0.06	0.04	0.13	0.15***
First release	−0.12*	−0.06	−0.36***	−0.44***
Adj.R^2	0.01	−0.02	0.10	0.11
W	3.69	2.11	13.17***	15.88***
		Total Investment		
Constant	0.03	0.17**	0.20	0.35
First release	−0.01	−0.04	−0.23***	−0.50***
Adj.R^2	−0.02	−0.01	0.16	0.25
W	0.15	5.04*	16.30***	25.00***
		Total Exports		
Constant	0.00	0.18	0.18	0.38
First release	−0.05	0.01	−0.08	−0.20
Adj.R^2	0.00	−0.03	0.00	0.03
W	0.81	1.48	2.14	2.83
		Total Imports		
Constant	0.20**	0.37***	0.15	0.59***
First release	−0.10***	−0.12***	−0.14*	−0.46***
Adj.R^2	0.09	0.23	0.08	0.39
W	11.60***	14.18***	4.23	29.50***

Note: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Newey-West standard errors with four lags. W is the wald test statistic for the hypothesis that the coefficients are jointly zero.

implying that the statistical agency uses all existing information efficiently to construct the data. Hence, adding further regressors to Equation (3.3) should not add much explanatory power. That is, for GDP releases to be perfectly efficient, all coefficients should remain jointly insignificant when we add other explanatory variables to the regression.

To test this, I augment the regressions with domestic and foreign macro variables. I obtain oil and stock prices from the OECD database[11] and the three month interbank rate, the consumer price index, the exchange rate, total employment, registered unemployed and Euroarea GDP from the real-time data set provided by Indergand and Leist (2014). The consumer sentiment index is taken from the SECO-Homepage.[12] Further results were obtained using additional variables such as government bond yields and other price data. However, I omit these (mostly insignificant) results for breavity.

Variables that are revised but only the latest vintage (or an incomplete real-time data set) is available should not be included in such an analysis. The reason is that correlations between variables may appear or disappear after the data has been revised but not using preliminary releases. But only the latter are available to the forecaster or the data producing agency. Furthermore, the results of including data that gets used by the SECO or SFSO to calculate QNA and ANA figures, such as employment or prices, should also be interpreted with caution. For example, if employment statistics help to explain future revisions, it could be that either national accounts data or employment data is produced inefficiently. Due to the limited amount of observations, each variable is included seperately.

Table 3.3 shows the regression results for GDP. In general, very little explanatory power is added by including additional variables. An exception are stock prices which both significantly predict revisions and add to the variation explained (captured by the $Adj.R^2$). The consumer sentiment index on

[11] OECD, "Main Economic Indicators - complete database", `http://dx.doi.org/10.1787/data-00052-en` (Accessed on July 12, 2016).

[12] `https://www.seco.admin.ch/seco/de/home/wirtschaftslage---wirtschaftspolitik/Wirtschaftslage/Konsumentenstimmung.html`. Note that the index was revised in $v = 2007q2$ and got revised back until $t = 1972q4$. For the time period before $t = 2007q2$ I therefore use the old official version.

the other hand is significant but hardly adds to the variation explained overall. A surprising result may be that preliminary releases of Euroarea GDP have a lot of explanatory power in the case of long term revisions, rev^{final}: Higher GDP growth abroad positively correlates with revisions to Swiss GDP. That only the final GDP revisions can be predicted may be related to the fact that benchmark revisions are coordinated across Europe. Often these methodological changes have similar effects on growth rates across countries and thus, revisions in different countries are correlated in the long term. However, this result also goes hand in hand with, and may be partly driven by the fact that revisions to total exports can also be forecasted by preliminary GDP releases abroad (see Table 3.E.4).

Table 3.E.1 - 3.E.5 show the results for the remaining QNA variables under consideration. Comparing these results to the findings in Table 3.2, the additional regressors add little explanatory power. There is some evidence that positive stock market returns are associated with positive revisions to private consumption, exports and imports and in all these cases, mean-reversion becomes even stronger. An appreciation of the exchange rate is associated with negative revisions to private consumption and negative final revisions to exports. Finally, both the consumer sentiment index and preliminary GDP releases for the Euroarea seem to be correlated with revisions to investment, exports and imports. However, these correlations materialize most strongly or exclusively in the long term and may be related to benchmark revisions as explained above. Some other variables are found to be significant (or jointly significant with the preliminary release) only for certain revision types but these instances should be interpreted with caution. Due to the limited amount of observations or extreme datapoints, a few spurious results are to be expected in these regressions.

Note that virtually all coefficients for the preliminary releases are negative, in many cases strongly and significantly, irrespective of the left hand side or the right hand side variable. If anything, mean-reversion becomes stronger if additional regressors are included. The longer the time horizon considered, the more negative and significant the coefficients of the preliminary releases

Table 3.3: Augmented Efficiency Regressions (OLS): GDP

	rev^{init}	rev^{ana-1}	rev^{ana}	rev^{final}
Constant	0.03	0.07**	0.00	0.12
First release	−0.02	−0.13***	0.02	−0.10
Oil Price	0.00*	0.00***	0.00	0.01**
Adj. R^2	−0.01	0.08	−0.04	0.09
W	0.93	10.01***	0.12	1.80
Constant	0.04	0.06	0.03	0.12
First release	−0.03	−0.10**	−0.05	−0.14
Stock Price	0.01***	0.01***	0.01***	0.03***
Adj. R^2	0.07	0.25	0.13	0.22
W	0.96	4.21	0.41	1.83
Constant	0.04	0.07*	0.02	0.06
First release	0.03	−0.05	0.08	0.08
Interest Rate	−0.03	−0.03	−0.05***	0.01
Adj. R^2	−0.01	−0.02	−0.02	−0.03
W	3.19	3.94	2.48	3.06
Constant	0.02	0.05	−0.01	0.07
First release	0.01	−0.04	0.03	0.01
CPI	0.00	−0.11	0.07	0.28**
Adj. R^2	−0.04	−0.01	−0.05	0.00
W	2.28	3.22	0.08	2.99
Constant	0.03	0.06*	0.01	0.13
First release	0.01	−0.05	0.04	0.03
Exchange Rate	−0.01	−0.02*	−0.01	−0.06***
Adj. R^2	−0.03	0.05	−0.05	0.07
W	1.90	3.16	1.08	7.02**
Constant	0.02	0.05	−0.02	0.09
First release	0.02	−0.08	0.04	0.14
Employment	−0.02	0.03	0.06	−0.21
Adj. R^2	−0.04	−0.04	−0.06	−0.02
W	0.42	1.82	0.16	3.21
Constant	0.05	0.08	0.03	0.13
First release	−0.05	−0.12	−0.02	−0.06
Unemployed	−0.01	−0.01	−0.01	−0.02
Adj. R^2	0.00	0.00	−0.02	0.00
W	2.05	4.02	0.56	3.76
Constant	0.11**	0.11***	0.16**	0.33**
First release	−0.12	−0.15*	−0.20	−0.30
Cons. Sentiment	0.00**	0.00	0.01***	0.01**
Adj. R^2	0.03	−0.01	0.10	0.07
W	4.40	12.35***	6.25**	10.41***
Constant	0.03	0.06**	0.00	0.14
First release	−0.05	−0.13***	−0.02	−0.44***
GDP Euroarea	0.07*	0.09	0.08	0.58***
Adj. R^2	−0.01	−0.01	−0.03	0.28
W	0.62	10.51***	0.08	8.51***

Note: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Newey-West standard errors with four lags. W is the wald test statistic for the hypothesis that the coefficients are jointly zero.

become. This confirms again that it takes at least a few quarters (vintages) for mean-reversion to materialize.

3.4.4 Sources of Data Inefficiency: Some Conjectures

The previous sections have shown that past revisions of national accounts variables in Switzerland can be characterized as mean reverting processes. The pattern is very consistent across variables, with GDP as an exception or borderline case, and becomes stronger when looking at longer term revisions. This essentially confirms what other authors have found for many economic variables abroad. Little explanatory power gets added by including additional variables in order to predict revisions. I find some indication that financial variables such as stock market returns or exchange rates may contain information that is not fully exploited by the statistical agencies. However, due to the limited amount of observations used in this study, more detailed research is necessary which is beyond the scope of this paper. In any case, mean reversion is clearly the dominant force for the inefficiency I find, with the exception of government consumption where both mean reversion and a negative bias are present. For the data providers as well as for future research, it is important to know the potential sources for these inefficiencies. In what follows, I provide some conjectures.

In principle and from a QNA perspective, there are at least four possible causes for mean reversion or biases.

First, it may be the case that the indicators employed for temporal disaggregation of the ANA data, are themselves inefficiently produced. To test this, a real-time data set for the underlying variables would be necessary. There is no such data set publicly available to date and hence, I can neither confirm nor rule out this possibility.

A second candidate is the approach employed for temporal disaggregation by the SECO. The inter- and extrapolation approaches, most notably Chow and Lin (1971), are best linear unbiased and they can also be formulated as a special case in a state space framework that can be estimated using the Kalman filter. The latter is known to produce optimal forecasts among linear estimators. Also given that one of the variables, GDP, is hardly affected by mean reversion (see Table 3.2), an inherent inefficiency of these employed methods is unlikely.[13]

The third candidate is seasonal adjustment. So far, almost all analyses of revisions in the literature have been based on seasonally adjusted data and there exists little evidence on the effect that these smoothing methods have on the properties of data revisions. A notable exception is Kavajecz and Collins (1995) who estimate Equation (3.3) for both seasonally adjusted and unadjusted monetary aggregates and find that seasonal adjustment may indeed introduce irrationality. More precisely seasonal adjustment may give rise to mean reverting preliminary releases. In Chapter 4 I investigate this question more thorougly based on simulated as well as real-world data and comes to the same conclusion.

Finally, any inefficiency inherent to ANA data will translate to the quarterly counterparts due to the dependency of the QNA on the annual benchmarks during the estimation process in Switzerland. Due to the short annual sample (see Section 2), meaningful statistical tests are hardly possible. The following section provides some graphical evidence on annual revisions.

3.4.5 A Closer Look at Annual Revisions

Figure 3.3 shows normalized revisions to annual growth rates for all six variables. All revisions have been divided by the standard deviation of the most recent (final) annual data of a variable in order to enhance comparability. The graphs on the left show the revision between the first estimate of annual growth rates, calculated by the SECO, and the second estimate, calculated by the SFSO. The first estimate (SECO) is based on quarterly extrapola-

[13] During earlier years, temporal disaggregation of an ANA aggregate was usually done by the SECO using two seperate steps. Disaggregation was carried out both using a level regression as well as a regression in first differences. The resulting quarterly series from the temporal disaggregation in levels was then used as an interpolated series. The resulting quarterly series from the regression in first differences was used for the extrapolation (growth rates from this series were used to extrapolate the interpolated levels). It is very hard to quantify the effects that this procedure may have on data revisions. However, the SECO abandonned this procedure completely in 2012 and thus should be irrelevant for the data of more recent years.

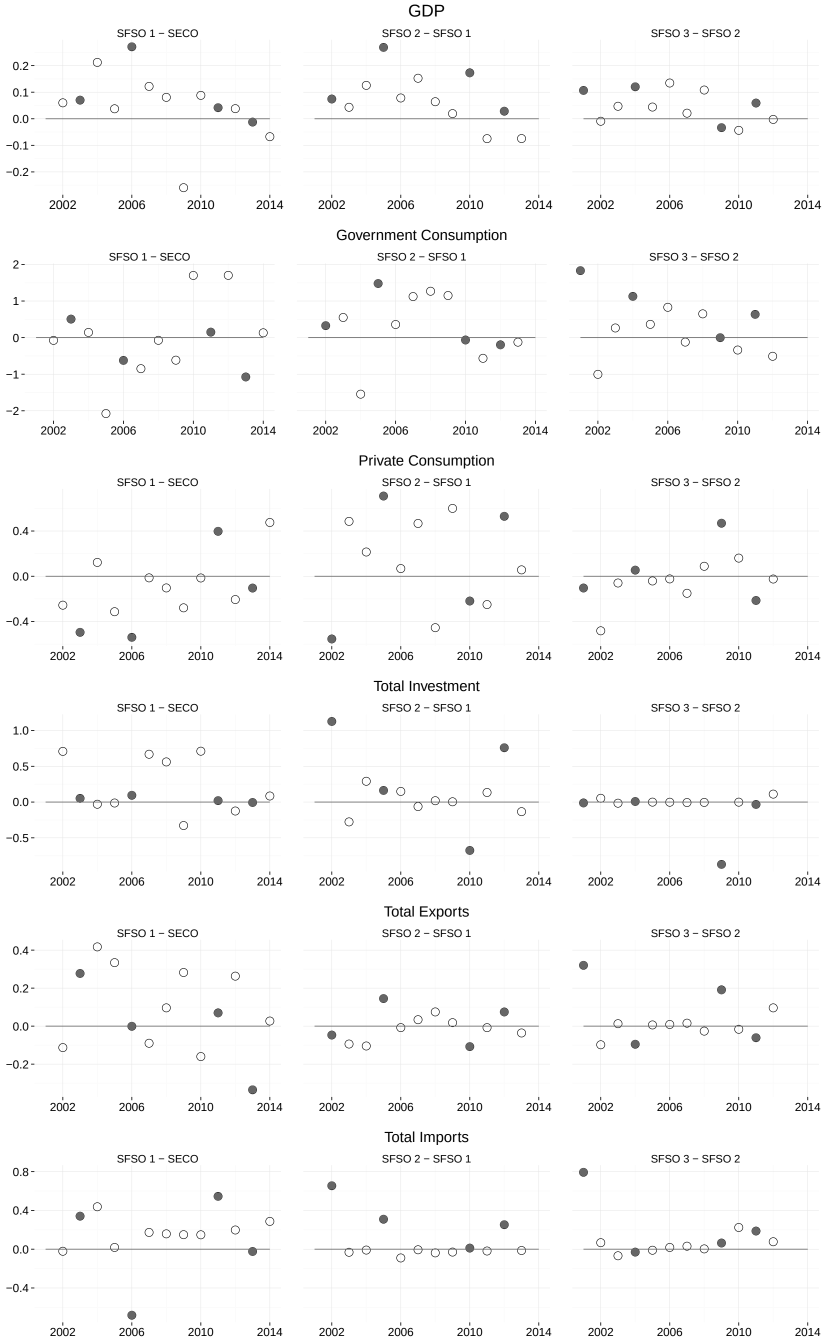
tions of the unrevised annual estimates of the SFSO and appears in the first quarter of any year. The second estimate (SFSO 1) uses a broader data basis and appears in the third quarter. The middle column of graphs compares the second with the third annual estimate (SFSO 2) appearing in the third quarter of the following year. The graphs on the right compare the second with the third annual estimate (SFSO 3) which incorporates the last revisions on an annual frequency apart from the more fundamental benchmark revisions. The filled dots correspond to annual revisions that have been affected by benchmark revisions.

From a visual inspection, revisions to annual GDP growth rates are smallest relative to the variability of the series (they are also smallest without normalization). On the other end, they are comparatively large for government consumption and private consumption, the two variables with relatively low variability. For the trade variables as well as total investment, the third annual revisions tend to be smaller than the second and the second annual revisions tend to be smaller than the first. For GDP, private and government consumption on the other hand, they shrink only slightly or, in fact, become bigger during the second annual revision.

With regard to the direction, a pattern of small positive revisions to annual GDP can be seen in all three graphs. This means that the annual figures calculated by the SECO slightly underestimated the first annual release by the SFSO (mean revision: 0.052%)[14] which in turn slightly underestimated later releases (mean revisions: 0.073% and 0.046%). This will of course translate also to the quarterly data to a certain extend (in particular, rev^{ana} and rev^{final} would be concerned above). However, these overall positive revisions are not large enough to prove statistically significant on the quarterly level where we have enough observations to perform meaningful statistical tests. Nevertheless, using average growth rates, Appendix 3 provides further evidence for a tendency of small positive revisions to GDP in the annual national accounts.

[14] mean revisions are also normalized by the standard deviation of the last vintage in order to enhance comparability across variables. Also note that I did not exclude benchmark revisions for these calculations due to the small sample.

Figure 3.3: Revisions to Annual Growth Rates



Note: The horizontal depicts the observation time for all graphs and the vertical the revision in percentage points. For example, the left datapoint in the first graph indicates that the SECO-release for $t = 2002$ was positively revised by the first SFSO-release. Left column: Revisions to the first annual GDP release. Middle column: Revisions to the second annual release. Right column: Revisions to the third annual release. Revisions have been normalized by the standard deviation of the most recent annual vintage. Annual benchmark revisions in dark.

On the other hand, it is notable that growth rates of government and private consumption during first annual revision (left column, mean revisions: -0.08% and -0.10%) were corrected downwards overall but then corrected upwards again with the second annual revision (middle column, mean revisions: 0.31% and 0.14%). On the other hand, for total imports, the first annual estimate was predominantly corrected upwards with the first annual revision but then corrected slighly downwards in the second. The only cases where annual revisions appear to be both relatively small and scattered on both sides of the zeroline in all three columns are total investment and total exports.

All this has to be treated with great caution as there are simply not enough observations on the annual level to reach definite conclusions. However these findings suggest that more research is warranted as soon as more (quarterly and annual) data vintages become available.

3.5 Conclusion

Revisions to Swiss national accounts data can be large. This paper updates and adds to the analyses of Cuche-Curti, Hall and Zanetti (2008), Siliverstovs (2011) and Bernhard (2016) for Swiss GDP and extends it to the most important expenditure-side components. For these components, revisions are often larger than for GDP. In particular, benchmark revisions typically concern the whole history of a variable, in some cases greatly affecting the attributes of the time series. For example, contractionary phases of Swiss GDP may not only shift in time but newly appear or disappear, at least if comparatively simple business cycle dating algorithms are used.

I test preliminary releases of all variables for inefficiency, emphasizing that one has to account for the effect of benchmark revisions. Albeit mean revisions for GDP, total investment, total exports and total imports are positive an average, I find no significant bias for these variables. That is, mean revisions are statistically not different from zero. For government consumption, however, first releases significantly overstated growth rates in the past and were consistently corrected downward lateron. Interestingly, in much later releases these estimates seem to converge back into the direction suggested by the preliminary quarterly extrapolations.

Moreover, I find for all GDP components mean reversion of initial releases. This means that extreme initial growth rates tend to be revised downward. This is very much in-line with the results of other studies that focused on national accounts data in other countries and also on different variables. Thus, mean reversion seems to be a predominant feature of preliminary announcements of seasonally adjusted economic data. Interestingly, at least in the case of the Swiss QNA, this pattern is hardly visible after only one revision and becomes much stronger in later releases.

Finally, I augment the forecast efficiency regressions with several other variables known at the time of the data release. While financial variables such as stock returns might contain some information not yet fully exploited, the variables considered in this paper generally add little predictive power. While this does not rule out that other information may help to predict revisions, it indicates that mean-reversion is the predominant force behind inefficient data releases.

The accuracy of initial releases is of utmost importance for policy makers. Any inefficiency in preliminary estimates should therefore be mitigated if possible. The fact that mean-reversion is internationally such a consistent feature and that it takes a considerable number of releases for it to materialize calls for further research.

Appendix

3.A Boxplots

Figure 3.A.1: Box plots for each vintage: Government Consumption

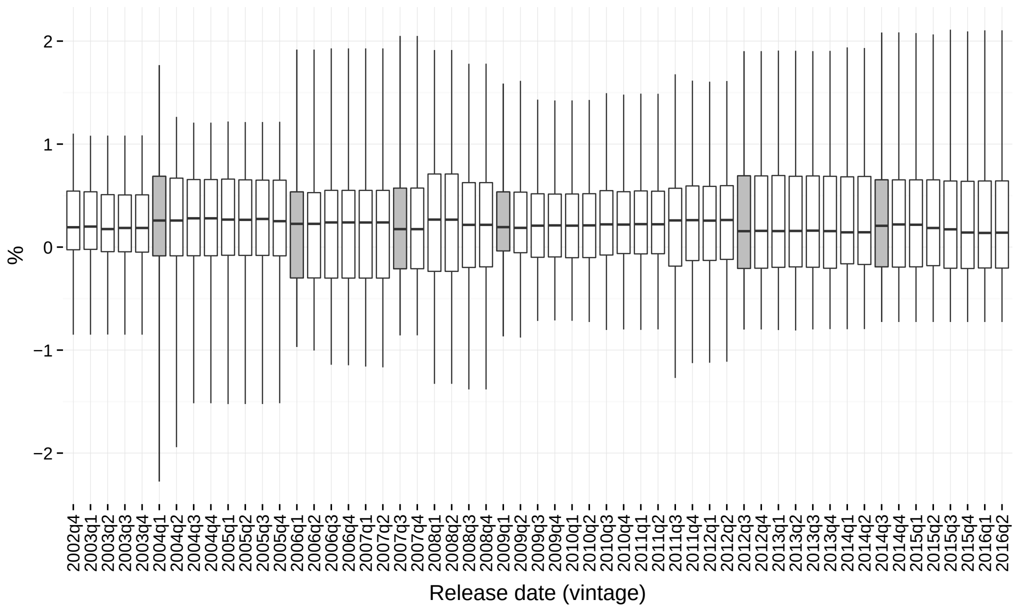


Figure 3.A.2: Box plots for each vintage: Private Consumption

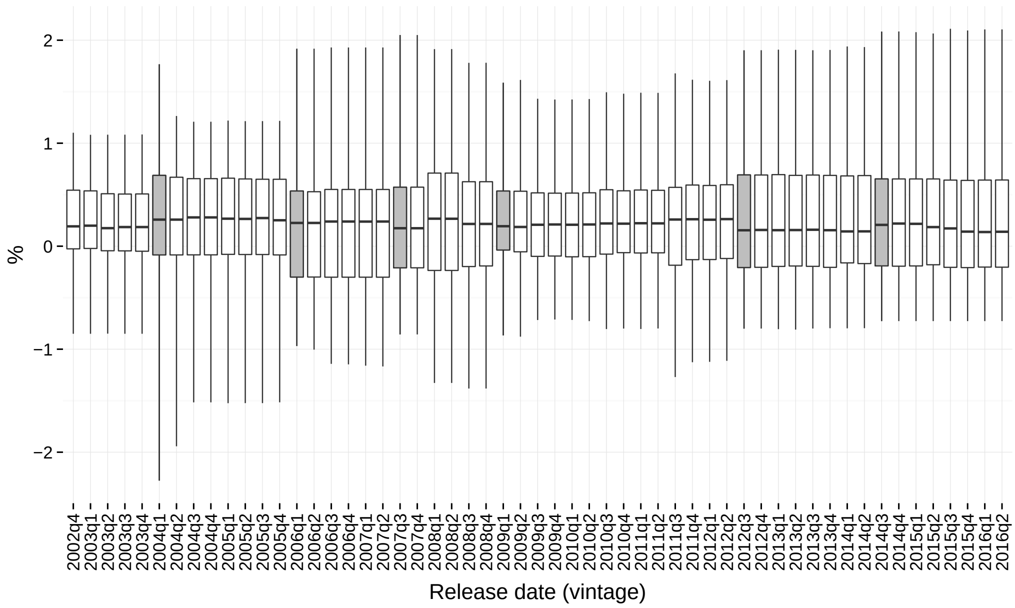


Figure 3.A.3: Box plots for each vintage: Total Investment

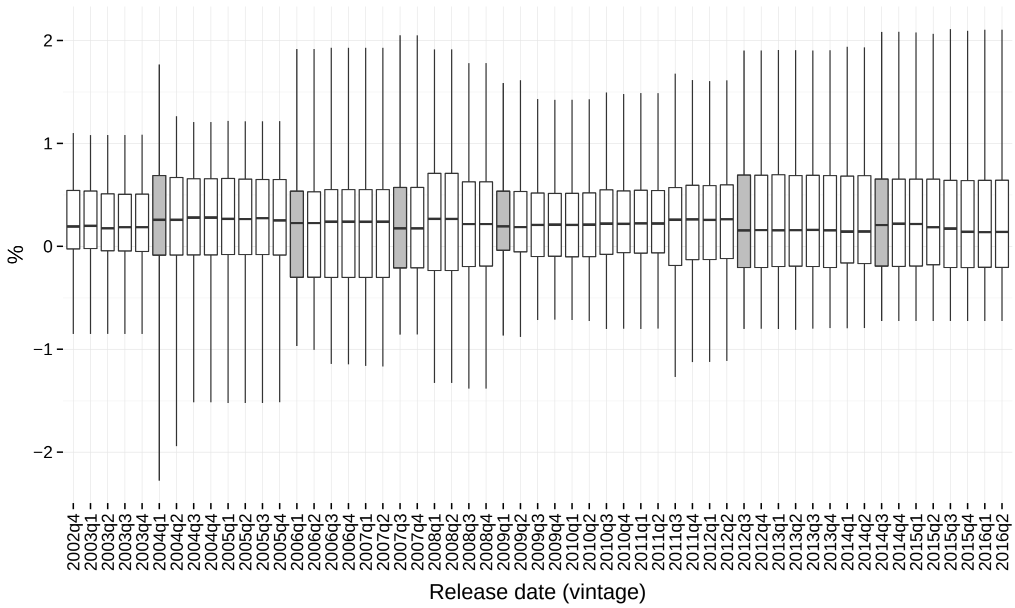


Figure 3.A.4: Box plots for each vintage: Total Exports

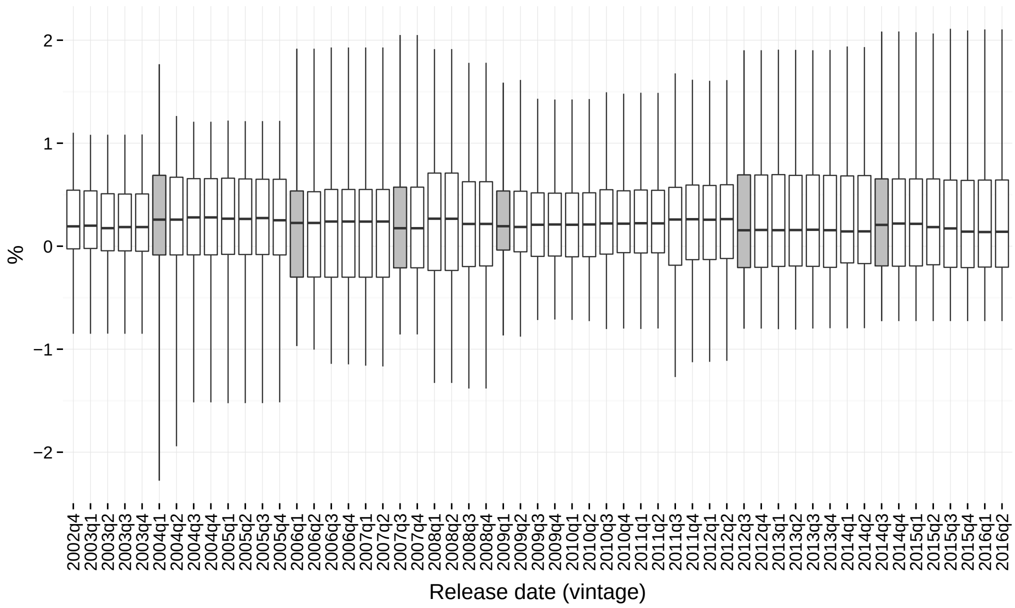
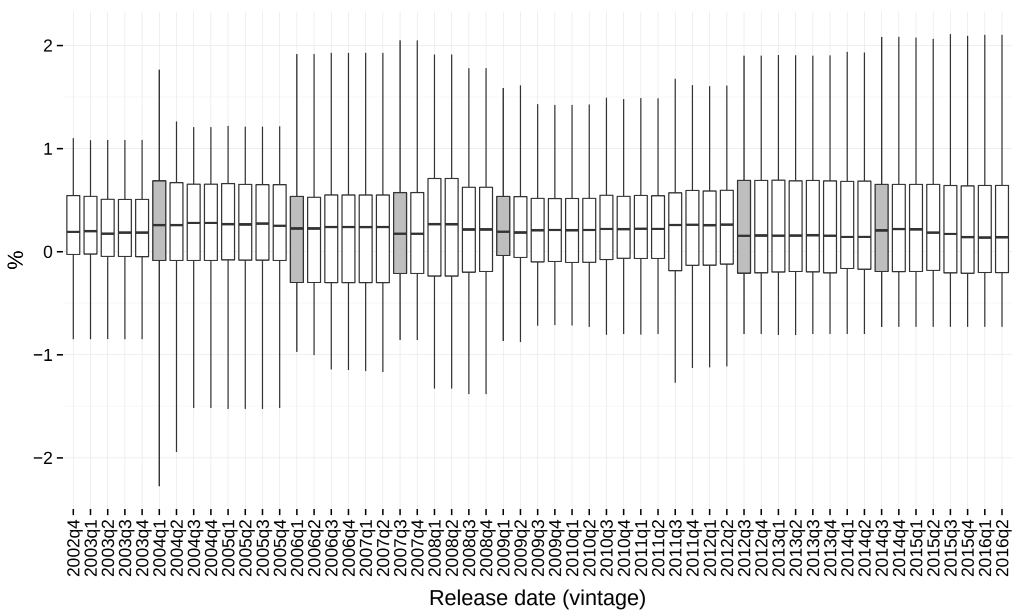


Figure 3.A.5: Box plots for each vintage: Total Imports



3.B GDP Revisions in Other Countries

Table 3.B.1 reproduces the results in Table 3.1 for GDP of other countries. For comparison the same vintages were used as in the Swiss case. Note however that these releases generally do not correspond to benchmark releases in these other countries. The results should hence be compared with care to Swiss GDP.

Table 3.B.1: Summary statistics for selected vintages

	2002q4	2004q1	2006q1	2007q3	2009q1	2012q3	2014q3
				Mean			
GDP US	0.70	0.72	0.71	0.71	0.71	0.74	0.75
GDP EA	0.49	0.50	0.52	0.52	0.52	0.52	0.52
GDP JP	0.33	0.35	0.30	0.31	0.31	0.29	0.29
				Standard deviation			
GDP US	0.58	0.56	0.55	0.55	0.55	0.57	0.56
GDP EA	0.43	0.44	0.44	0.43	0.43	0.43	0.43
GDP JP	1.00	0.79	0.79	0.82	0.81	0.91	0.90
				First order autocorrelation			
GDP US	0.35	0.33	0.32	0.32	0.32	0.27	0.29
GDP EA	0.42	0.38	0.43	0.45	0.43	0.41	0.41
GDP JP	-0.08	0.21	0.15	0.13	0.15	-0.05	-0.04
				Number of negative growth rates			
GDP US	7	6	5	5	5	5	4
GDP EA	5	6	4	4	4	4	5
GDP JP	15	16	16	16	16	19	20

Note: All calculations based on the period t=1990q1-2002q3

3.C Revisions to mean annual GDP growth

Figure 3.C.1 shows the revisions to the mean of annual growth rates. For each vintage, the most recent 10 years were used for computation:

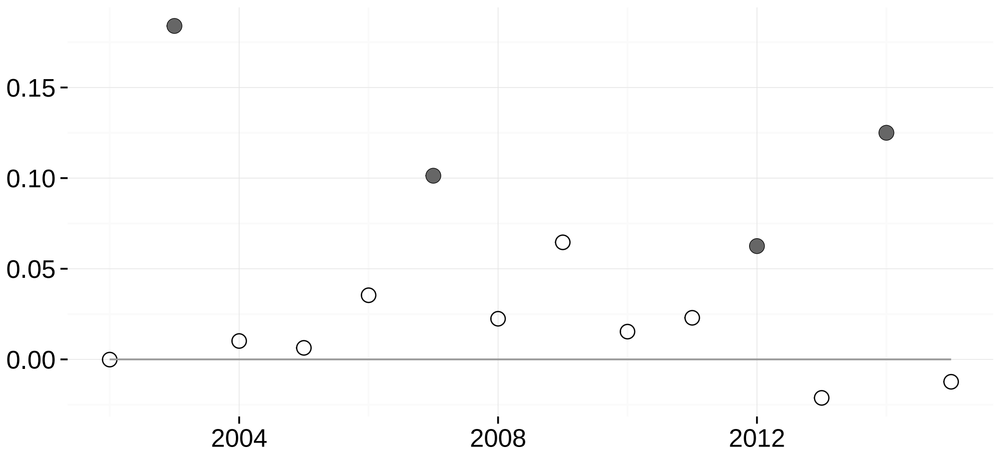
$$\Delta\overline{gdp}^{\,v} = \frac{1}{10}\sum_{t=T^v-9}^{T^v} \Delta gdp_t^v \tag{3.11}$$

where the frequency of the data now is annual and T^v stands for the last observation in any given vintage, v. Revisions to mean growth are then defined as

$$rev^{ana} = \overline{gdp}^{\,ana} - \overline{gdp}^{\,ana-1} \tag{3.12}$$

where ana stands for all vintages at which the ANA data got revised or newly released. Hence, contrary to the statistics in Table 3.1 which are all computed over the same horizon, $t = 1990q1 - 2002q3$, the period used for computation varies for each revision depicted in Figure 3.C.1. Despite the fact that there are very few observations, revisions clearly seem to be positively skewed indicating a tendency for underestimation of growth rates. However, note that Figure 3.C.1 also includes benchmark revisions and, for further investigation, one would have to exclude those (leaving us with even less observations) or take into account other countries for comparison.

Figure 3.C.1: Revisions to annual average growth of GDP



Note: The figure depicts the annual revisions to annual average growth. A positive observation indicates that the mean growth of the 10 years prior to T was revised upwards. Benchmark revisions in dark.

3.D Preliminary Releases and their Revisions

Figure 3.D.1: First Estimates and Revisions, Private Consumption

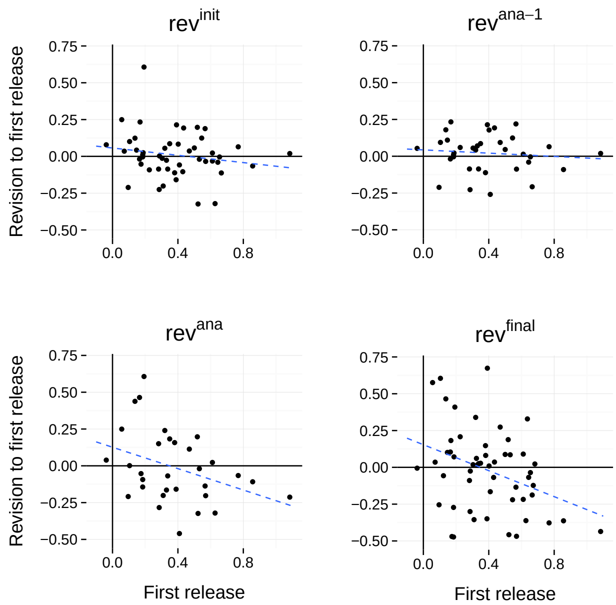


Figure 3.D.2: First Estimates and Revisions, Total Investment

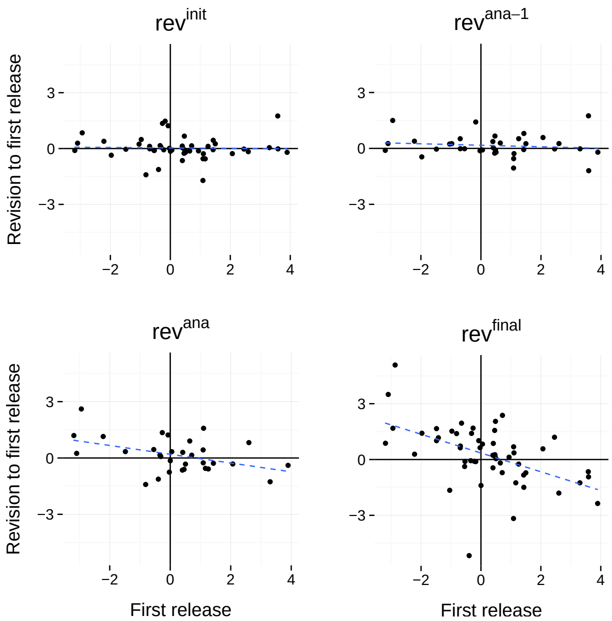


Figure 3.D.3: First Estimates and Revisions, Total Exports

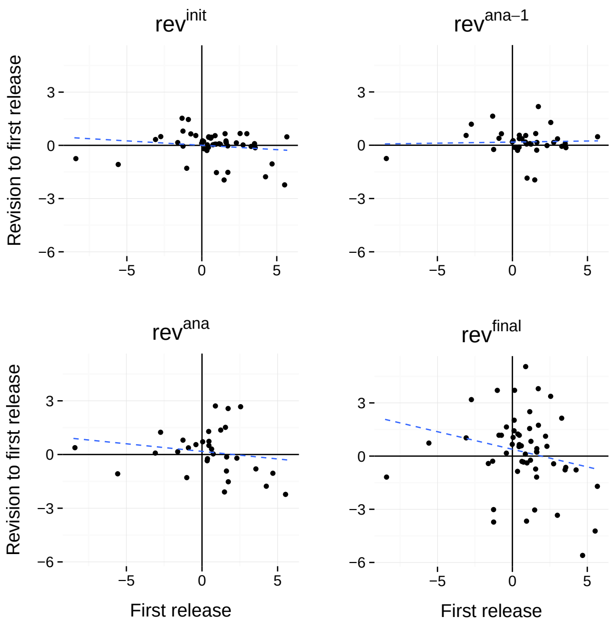
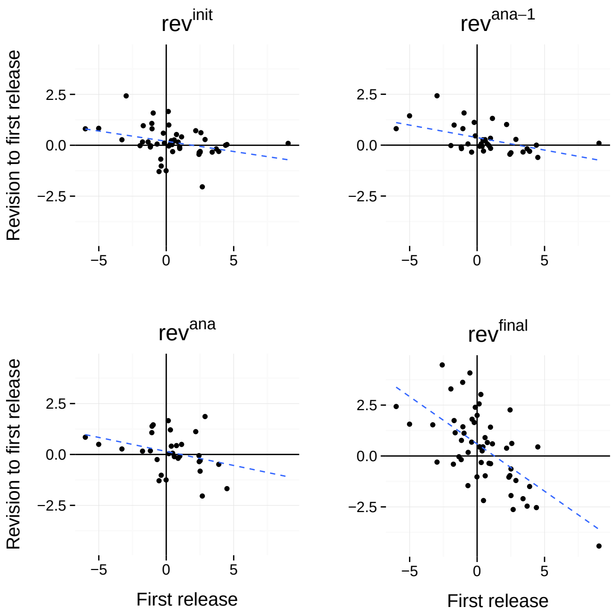


Figure 3.D.4: First Estimates and Revisions, Total Imports



3.E Tables

Table 3.E.1: Augmented Efficiency Regressions (OLS): Private Consumption

	rev^{init}	rev^{ana-1}	rev^{ana}	rev^{final}
Constant	0.06	0.05**	0.14*	0.16***
First release	−0.15*	−0.08**	−0.40***	−0.46***
Oil Price	0.00	0.00	0.00	0.00
R^2	0.02	−0.01	0.10	0.10
W	3.57	7.25**	12.42***	13.31***
Constant	0.10***	0.08***	0.20***	0.20***
First release	−0.24***	−0.14***	−0.54***	−0.57***
Stock Price	0.01***	0.01***	0.01***	0.01***
R^2	0.16	0.20	0.30	0.16
W	10.34***	13.58***	14.17***	21.58***
Constant	0.06	0.06**	0.13	0.16***
First release	−0.12	−0.04	−0.36***	−0.43***
Interest Rate	−0.01	−0.03	0.00	−0.01
R^2	0.00	−0.01	0.07	0.10
W	3.38	5.35*	8.10**	13.37***
Constant	0.06	0.05	0.12	0.17***
First release	−0.12*	−0.06	−0.36***	−0.45***
CPI	0.00	−0.05	0.02	−0.16
R^2	−0.01	−0.04	0.07	0.13
W	3.80	2.01	9.72***	15.48***
Constant	0.09**	0.07***	0.18**	0.21***
First release	−0.16**	−0.09**	−0.39***	−0.51***
Exchange Rate	−0.03***	−0.02**	−0.04*	−0.04***
R^2	0.17	0.05	0.17	0.23
W	6.36**	8.89***	10.58***	22.43***
Constant	0.07*	0.05*	0.11	0.14***
First release	−0.10*	−0.05	−0.39***	−0.48***
Employment	−0.12*	−0.04	0.16*	0.15*
R^2	0.04	−0.04	0.12	0.12
W	3.92	3.24	15.23***	18.50***
Constant	0.05	0.04	0.14*	0.17***
First release	−0.12*	−0.06	−0.36***	−0.45***
Unemployed	0.00	0.00	−0.01**	−0.01*
R^2	−0.01	−0.05	0.13	0.11
W	3.52	2.18	9.43***	15.80***
Constant	0.06	0.06**	0.20***	0.17***
First release	−0.13**	−0.07	−0.43***	−0.46***
Cons. Sentiment	0.00	0.00	0.00***	0.00
R^2	−0.01	−0.04	0.18	0.10
W	4.20	6.45**	9.48***	15.06***
Constant	0.06	0.04	0.13	0.16***
First release	−0.12*	−0.06	−0.36***	−0.40***
GDP Euroarea	−0.02	0.04	−0.01	−0.11**
R^2	0.00	−0.03	0.07	0.14
W	3.04	1.74	10.50***	11.33***

Note: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. Newey-West standard errors with four lags. W is the wald test statistic for the hypothesis that the coefficients are jointly zero.

Table 3.E.2: Augmented Efficiency Regressions (OLS): Government Consumption

	rev^{init}	rev^{ana-1}	rev^{ana}	rev^{final}
Constant	0.03	−0.08***	−0.11	0.00
First release	−0.29***	−0.11	−0.36***	−0.31**
Oil Price	0.00	0.00*	−0.01	−0.01***
R^2	0.19	0.10	0.20	0.10
W	18.94***	28.05***	42.11***	6.72**
Constant	0.02	−0.08**	−0.14	−0.02
First release	−0.28***	−0.11	−0.32**	−0.29*
Stock Price	0.01	0.01*	0.00	−0.01
R^2	0.19	0.09	0.12	0.05
W	18.91***	31.05***	20.36***	5.96**
Constant	0.05	−0.09	0.06	0.08
First release	−0.30***	−0.10	−0.39***	−0.32**
Interest Rate	−0.04	0.01	−0.24**	−0.14
R^2	0.19	0.05	0.26	0.07
W	20.04***	17.93***	18.87***	4.56
Constant	0.02	−0.10**	−0.07	0.05
First release	−0.29***	−0.10	−0.39***	−0.36**
CPI	0.01	0.09	−0.35	−0.60
R^2	0.18	0.06	0.16	0.09
W	21.63***	19.36***	26.82***	5.88**
Constant	0.02	−0.08***	−0.18	−0.02
First release	−0.29***	−0.10	−0.31**	−0.28*
Exchange Rate	0.01	−0.01	0.03	−0.01
R^2	0.19	0.06	0.13	0.04
W	17.73***	15.10***	23.15***	6.33**
Constant	0.08	−0.04	−0.07	0.12
First release	−0.29***	−0.10	−0.35***	−0.27**
Employment	−0.30	−0.23**	−0.42	−0.79***
R^2	0.22	0.14	0.17	0.13
W	21.23***	15.77***	15.65***	5.39*
Constant	0.02	−0.08***	−0.14	−0.03
First release	−0.32***	−0.13*	−0.39***	−0.32**
Unemployed	0.02***	0.01**	0.02	0.02
R^2	0.24	0.13	0.15	0.06
W	29.86***	34.08***	22.37***	6.95**
Constant	0.02	−0.09***	−0.16	−0.04
First release	−0.29***	−0.13	−0.37***	−0.30
Cons. Sentiment	0.00	0.00	0.00	0.00
R^2	0.18	0.07	0.14	0.04
W	14.20***	41.58***	20.91***	3.12
Constant	0.02	−0.05	−0.14	0.04
First release	−0.29***	−0.13	−0.32**	−0.33**
GDP Euroarea	−0.01	−0.11	−0.01	−0.22
R^2	0.18	0.08	0.12	0.06
W	22.52***	9.72***	24.63***	5.77*

Note: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Newey-West standard errors with four lags. W is the wald test statistic for the hypothesis that the coefficients are jointly zero.

Table 3.E.3: Augmented Efficiency Regressions (OLS): Total Investment

	rev^{init}	rev^{ana-1}	rev^{ana}	rev^{final}
Constant	0.03	0.17**	0.19	0.34*
First release	−0.02	−0.04	−0.28***	−0.56***
Oil Price	0.00	0.00	0.01*	0.02***
R^2	−0.04	−0.04	0.17	0.28
W	0.25	5.09*	12.63***	37.38***
Constant	0.03	0.16*	0.19	0.33
First release	−0.02	0.00	−0.21**	−0.58***
Stock Price	0.00	−0.02	−0.01	0.04
R^2	−0.04	−0.02	0.13	0.27
W	0.22	3.76	6.53**	31.35***
Constant	0.01	0.19*	0.14	0.28
First release	−0.01	−0.04	−0.22***	−0.49***
Interest Rate	0.03	−0.02	0.08	0.10
R^2	−0.04	−0.05	0.13	0.24
W	0.04	3.90	10.96***	24.84***
Constant	0.01	0.16	0.18	0.31
First release	−0.01	−0.04	−0.23***	−0.50***
CPI	0.30	0.14	0.22	0.69
R^2	−0.02	−0.04	0.13	0.25
W	0.08	2.87	16.50***	22.86***
Constant	0.07	0.18**	0.25*	0.41*
First release	−0.02	−0.05	−0.24***	−0.52***
Exchange Rate	−0.04	−0.02	−0.04	−0.09*
R^2	−0.02	−0.04	0.14	0.25
W	0.60	4.58	13.78***	29.56***
Constant	0.03	0.20*	0.22	0.26
First release	−0.01	−0.04	−0.23***	−0.50***
Employment	−0.01	−0.14	−0.14	0.45
R^2	−0.04	−0.04	0.13	0.24
W	0.11	4.01	16.44***	24.16***
Constant	0.06	0.17**	0.22*	0.45***
First release	−0.02	−0.04	−0.24***	−0.54***
Unemployed	−0.02	0.00	−0.01	−0.10***
R^2	−0.01	−0.05	0.13	0.35
W	0.66	5.70*	17.55***	35.02***
Constant	0.08	0.20***	0.13	0.71***
First release	−0.03	−0.05	−0.21***	−0.61***
Cons. Sentiment	0.00	0.00	−0.01	0.03*
R^2	−0.03	−0.04	0.14	0.35
W	0.40	10.06***	8.18**	62.81***
Constant	−0.03	0.12	0.15	0.09
First release	−0.05	−0.06	−0.27***	−0.62***
GDP Euroarea	0.37***	0.28*	0.38***	1.45***
R^2	0.07	−0.01	0.21	0.48
W	0.66	4.27	9.99***	51.54***

Note: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Newey-West standard errors with four lags. W is the wald test statistic for the hypothesis that the coefficients are jointly zero.

Table 3.E.4: Augmented Efficiency Regressions (OLS): Total Exports

	rev^{init}	rev^{ana-1}	rev^{ana}	rev^{final}
Constant	0.02	0.20	0.18	0.42*
First release	−0.07	−0.04	−0.13*	−0.29***
Oil Price	0.01	0.02**	0.01	0.03*
R^2	0.00	0.04	−0.02	0.06
W	1.21	3.09	5.04*	7.47**
Constant	0.02	0.21	0.23	0.42**
First release	−0.08	−0.04	−0.17***	−0.36***
Stock Price	0.02	0.04***	0.06***	0.13***
R^2	0.01	0.06	0.10	0.18
W	1.55	3.25	11.39***	9.30***
Constant	0.06	0.29	0.32	0.47*
First release	−0.05	0.01	−0.09	−0.20
Interest Rate	−0.10	−0.15	−0.20*	−0.14
R^2	−0.01	−0.03	−0.02	0.02
W	0.38	1.74	4.33	4.41
Constant	0.01	0.22	0.21	0.42
First release	−0.04	0.03	−0.06	−0.14
CPI	−0.27	−0.55	−0.43	−1.43*
R^2	−0.01	−0.02	−0.03	0.06
W	0.35	1.98	2.32	2.79
Constant	0.01	0.16	0.19	0.65**
First release	−0.05	0.02	−0.09	−0.28**
Exchange Rate	0.00	0.03	−0.01	−0.31***
R^2	−0.02	−0.06	−0.04	0.12
W	0.81	1.62	1.96	5.60*
Constant	−0.06	0.14	0.14	0.48
First release	−0.05	0.01	−0.08	−0.20
Employment	0.35	0.19	0.26	−0.49
R^2	0.00	−0.05	−0.03	0.02
W	1.49	0.51	1.90	3.82
Constant	0.02	0.19	0.19	0.37
First release	−0.05	0.01	−0.09	−0.20
Unemployed	−0.01	−0.01	−0.01	0.01
R^2	−0.02	−0.05	−0.03	0.01
W	0.70	1.90	2.04	2.03
Constant	0.03	0.25	0.51***	0.64**
First release	−0.06	0.00	−0.19**	−0.26*
Cons. Sentiment	0.00	0.01	0.03**	0.02
R^2	−0.02	−0.04	0.09	0.03
W	0.59	1.71	11.23***	4.81*
Constant	−0.03	0.10	0.12	0.27
First release	−0.10	−0.05	−0.21***	−0.36**
GDP Euroarea	0.36	0.62**	0.72***	1.19**
R^2	0.03	0.04	0.04	0.08
W	2.87	2.11	16.08***	6.97**

extitNote: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Newey-West standard errors with four lags. W is the wald test statistic for the hypothesis that the coefficients are jointly zero.

Table 3.E.5: Augmented Efficiency Regressions (OLS): Total Imports

	rev^{init}	rev^{ana-1}	rev^{ana}	rev^{final}
Constant	0.20**	0.37***	0.14	0.58***
First release	−0.10***	−0.12**	−0.15**	−0.48***
Oil Price	0.00	0.00	0.01	0.02*
R^2	0.07	0.21	0.06	0.39
W	11.11***	13.79***	4.97*	38.50***
Constant	0.20**	0.38***	0.18	0.59***
First release	−0.12***	−0.14***	−0.19***	−0.48***
Stock Price	0.02	0.02	0.04***	0.02
R^2	0.13	0.26	0.17	0.38
W	12.77***	15.58***	9.99***	32.14***
Constant	0.21*	0.46***	0.12	0.70***
First release	−0.10***	−0.12**	−0.13*	−0.47***
Interest Rate	−0.03	−0.13**	0.04	−0.16
R^2	0.08	0.23	0.05	0.38
W	10.66***	12.77***	3.36	39.52***
Constant	0.18*	0.38***	0.08	0.59***
First release	−0.10***	−0.12**	−0.15**	−0.46***
CPI	0.22	−0.11	0.67**	0.02
R^2	0.08	0.21	0.10	0.37
W	10.67***	15.36***	5.30*	28.43***
Constant	0.25***	0.34***	0.29*	0.67***
First release	−0.11***	−0.12***	−0.15**	−0.47***
Exchange Rate	−0.06	0.06	−0.12	−0.12
R^2	0.12	0.24	0.11	0.40
W	12.04***	11.23***	7.33**	33.60***
Constant	0.17*	0.46***	0.10	0.68***
First release	−0.10***	−0.11**	−0.14*	−0.46***
Employment	0.15	−0.48*	0.32	−0.48
R^2	0.08	0.25	0.06	0.38
W	10.32***	21.14***	3.91	29.77***
Constant	0.23**	0.36***	0.21	0.66***
First release	−0.11***	−0.12**	−0.16**	−0.49***
Unemployed	−0.02	0.01	−0.03	−0.06***
R^2	0.10	0.21	0.08	0.41
W	11.04***	12.93***	7.64**	45.82***
Constant	0.28**	0.30***	0.35***	0.83***
First release	−0.12***	−0.10**	−0.20***	−0.52***
Cons. Sentiment	0.01	−0.01	0.02***	0.02**
R^2	0.09	0.22	0.14	0.41
W	7.99**	7.66**	19.02***	54.60***
Constant	0.14*	0.33***	0.10	0.46***
First release	−0.13***	−0.15**	−0.18***	−0.52***
GDP Euroarea	0.40***	0.27	0.39***	0.83***
R^2	0.16	0.23	0.12	0.43
W	11.57***	16.80***	7.72**	51.51***

extitNote: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Newey-West standard errors with four lags. W is the wald test statistic for the hypothesis that the coefficients are jointly zero.

Chapter 4

Seasonal Adjustment and Data Inefficiency: Evidence from Simulation and Real-World Data

Ronald Indergand

Summary: Seasonal adjustment based on moving-average filters is suboptimal from a forecast efficiency perspective. First, subsequent revisions to the data are comparatively large. Second, initial announcements are of mean-reverting character, that is, large initial growth rates tend to be revised downward. Both features increase the difficulty of assessing the dynamics of a variable and lead to suboptimal signals for policy makers. This paper shows by simulation that model-based seasonal adjustment reduces the first feature of the data and eliminates the second if the data generating process is known. This results in more efficient preliminary estimates allowing for a more accurate assessment of the current state of an economic variable. Using GDP data from nine countries I demonstrate in a real-world setting how seasonal adjustment produces mean-reverting preliminary GDP releases. Overall, seasonal adjustment may account for the bulk of the results regarding mean-reverting data revisions that have been found by numerous studies.

I thank William R. Bell, Francis X. Diebold, Daniel Kaufmann, Isabel Martinez, Brian Monsell, Klaus Neusser, David Romer, Ulrich Woitek and Jonathan Wright for valuable comments and suggestions. I also gratefully acknowledge financial support from the Swiss National Science Foundation for this project.

4.1 Introduction

Seasonal variation in economic data related to the weather, holidays and other sources obscure movements related to the business cycle. It is therefore the seasonally adjusted (SA) series that matters for the assessment of the current economic conditions and hence for economic policy. By its nature, however, seasonal adjustment introduces subsequent revisions to the data. These are often large even if the not seasonally adjusted (NSA) counterpart remains the same for all observations. For example, US employment from the household survey as well as registered unemployed in Switzerland do not get revised after they have been published. Nevertheless, revisions in the seasonally adjusted data on the other hand are substantial. In the case of Swiss unemployment, revisions to the preliminary unemployment change are on average and in absolute terms about half as big as the preliminary change itself. In fact, of all observations about 18% in the case of Swiss registered unemployed and about 9% in the case of US employment (household survey) indicated a change of the series in the wrong direction when they were first released. That is, a rise of Swiss unemployment was indicated instead of a fall or vice versa in more than one out of six data releases, and in almost one out of ten data releases in the case of US employment.[1] This finding may be particularly striking in the light of the fact that (un)employment figures belong to the most closely watched economic figures both by policy makers and by the public.

Furthermore, the preliminary releases in both examples are predictable and therefore suboptimal. The problem originates from the specific seasonal filter employed. By simulation I show that model-based filters deliver superior results in this regard compared to moving-average-based filters at least if the data generating process (DGP) is known and the default settings of X-13ARIMA-SEATS are used. X-13ARIMA-SEATS is the Census Bureau's newest software which allows for both types of adjustment. The first, parametric approach (SEATS) uses the information of the (estimated)

[1] In both examples, revisions are defined as the difference between first releases and the final vintage available at the time of writing. Changes are month-over-month first differences.

DGP explicitly in order to model unobserved components of the series. A Wiener-Kolmogorov type filter is then used for seasonal adjustment. The second approach belongs to the so-called X-11 family and corresponds to the algorithm implemented in the Census Bureau's X-12-ARIMA program (throughout the paper, I use the term "X-11" to refer to the implementation of moving average based filters in X-13ARIMA-SEATS). Contrary to model-based adjustment, X-11 does not explicitly use the estimated model of the DGP for the seasonal decomposition.

Data producing agencies face a myriad of decisions and challenges when they adjust series for seasonal effects. These include for example the parameter choice when fitting an ARIMA model, outlier and calender effects detection, or the fundamental issue of choosing between different adjustment approaches. To make matters worse, we can generally never observe the true nonseasonal series nor will we ever know its data generating process. This makes it very hard to distinguish 'good' from 'bad' adjustment. All this makes seasonal adjustment an equally difficult as well as important task. It should be carried out very carefully and its effects should be well understood. Surprisingly, however, the literature has essentially turned its back on issues related to seasonality during recent years (two notable exceptions are Wright, 2013 and Boldin and Wright, 2015). There exists a considerable body of literature dating back to the time when the X-11 family of seasonal filters was developed and widely adopted by major statistical offices during the 1970s and 1980s. It is well known in this literature that different seasonal adjustment techniques may result in very different estimates of the smoothed time series. For example, Maravall (1980) shows that two alternative seasonal adjustment methods result in different estimates of monetary aggregates, and as a consequence monetary policy decisions would differ with a probability of more than 10%. In fact, the German statistical agency Destatis has adopted the practice of publishing different versions of seasonally adjusted GDP, using both the X-12-ARIMA and the BV4.1 procedure (Speth, 2004).

Several authors have suggested to use revisions to the adjusted series, or the seasonal components respectively, as a criteria for choosing between different seasonal adjustment methods (see for example Pierce, 1980 or Dagum and Morry, 1984). Dagum and Laniel (1987) show that revisions are better behaved using X-11-ARIMA adjustment as opposed to the original X-11 algorithm. More recently, Kavajecz and Collins (1995) find that preliminary releases may become inefficient using X-11-ARIMA, in particular if non-concurrent adjustment is used.[2] Revision statistics should certainly not be the single most important, let alone the only criteria for good seasonal adjustment. Bell and Hillmer (1984) for example, argue that the decision should rather be based on information in the data, beliefs about seasonality and the objects of seasonal adjustment. Nevertheless, revision statistics provide important guidance on how stable the estimates are. Furthermore, provided that different seasonal filters converge to similar final values, they may be used as an important quality indicator for choosing among different seasonal adjustment techniques.

Moreover, data revisions in general have once again become a focus of the recent literature. Studies have found that revisions to national accounts data (such as GDP) are not optimal in several respects. First, revisions are in some cases biased, that is, they do not have a zero mean, and their variance is found to be large (Aruoba, 2008). Second, many authors find that revisions can be described as reducing noise as opposed to adding news, indicating that initial announcements are noisy, suboptimal estimates of the revised value (see the literature overview in Jacobs and Van Norden, 2011). In particular, initial announcements are in most cases found to exhibit a strong mean-reverting tendency. That is, high initial growth rates tend to be revised downward and low initial growth rates tend to be revised upward in later releases. For example, Mankiw, Runkle and Shapiro (1984) find mean-reversion for the money stock in the US, Faust, Rogers and Wright (2005) find it for preliminary GDP releases of most G7 countries, Garratt and Vahey (2006) find it for all expenditure side GDP components, the monetary base, industrial production, unemployment, retail sales and other variables in the UK and in Chapter 3 I found mean-reversion also for all expenditure side

[2] In concurrent adjustment, the seasonal effects are newly estimated whenever a new observation appears. In non-concurrent adjustment, the seasonal effects are kept constant and usually revised only once a year.

GDP components in Switzerland. All mentioned studies focuse on seasonally adjusted data.

The results in this paper suggest that studying seasonal adjustment helps both to explain these findings and to mitigate the problems that arise from ill-behaved revisions. I show that data revisions depend heavily on the adjustment method. The magnitude of revisions differs between the X-11 approach and SEATS even if the data generating process (DGP) is known and the models are correctly specified. Furthermore, I find that moving average filters generate preliminary releases with a strong mean-reverting tendency. I show by simulation that this seems to be a general feature of moving average type filters at least for certain DGPs. In-line with the findings in Faust, Rogers and Wright (2005) and in Chapter 3 of this dissertation, the pattern I find for X-11 generated revisions becomes stronger looking at long-term revisions as opposed to short-term revisions and the coefficients are of similar magnitude.

As seasonal adjustment methods were changed frequently in practice, it is impossible to exactly quantify what their effect was on total revisions of real-world data. However, seasonally adjusting GDP data from nine countries in a pseudo real-time setting indicates that seasonal adjustment can give rise to mean-reverting initial releases up to the extent found by Faust, Rogers and Wright (2005). Most of these countries used to or are still employing filters from the X-11 family. The adoption of a model-based seasonal adjustment technique may result in more efficient estimates, albeit the results deteriorate if model uncertainty is introduced. Along the lines of Wright (2013) I therefore advocate employing filters that use a model-based decomposition. Revisions are both smaller and the mean-reverting tendency gets reduced, such that the resulting statistics gain in reliability and false signals for policy makers become less frequent.

The remainder of this paper is organized as follows. Using two examples, Section 2 shows that seasonal adjustment causes sizeable revisions and gives rise to mean-reverting preliminary releases. Section 3 provides some background on seasonal adjustment and briefly describes the two main approaches that are used in practice. Section 4 shows by simulation, that these two approaches exhibit different properties with respect to data revisions and that X-11 type adjustment gives rise to mean-reverting preliminary estimates. Building on this finding, Section 5 suggests that alternative seasonal adjustment methods could improve the properties of preliminary releases of official statistics such as unemployment and GDP figures. Section 6 concludes.

4.2 A Real-World Example

Prominent examples where revisions are solely due to seasonal adjustment can be found in labour market statistics. The Bureau of Labor Statistics collects US employment both according to a household survey and a payroll survey. Unlike payroll employment, the rawdata for employment from the household survey is not revised after it has been published. For the considered time period, the BLS used the X-12-ARIMA software for seasonal adjustment.[3] The adjustment is carried out on a desaggregated level (the components are adjusted and then summed to form total employment) and since 2004 using a concurrent approach. That is, the seasonal factors are updated for the current month's estimate, however the previous months' estimates are not revised until the end of the calendar year. For growth rates this produces a mixture between concurrent and non-concurrent adjustment.

On the other hand, Swiss registered unemployment data collected by the State Secretariat for Economic Affairs (SECO) is directly adjusted on the aggregated level and a concurrent approach is used. That is, the seasonal factors are updated with every release for the whole time series. Seasonal adjustment is performed using Demetra+, a program developed by Eurostat which implements the US Census Bureau's X-12-ARIMA algorithm (see Section 3).[4]

[3] The BLS switched from using X-11-ARIMA to X-12-ARIMA in 2003 and recently switched to the new X-13ARIMA-SEATS software in 2015. However, the BLS still uses the X-11 type adjustment embedded in the X-13 software which corresponds to X-12-ARIMA. For more details see Tiller and Evans (2016). Vintage data can be obtained from the OECD real-time database.

[4] The SECO publishes only the latest two seasonally adjusted figures in its monthly bulletin *Die Lage am Arbeitsmarkt*. I obtained these official releases in real-time from the bulletin starting in 2003m6. Longer seasonally adjusted time series are not regularly published but can be obtained from the SECO. Furthermore, the SECO maintains

Figure 4.1: Unemployment Level in Switzerland

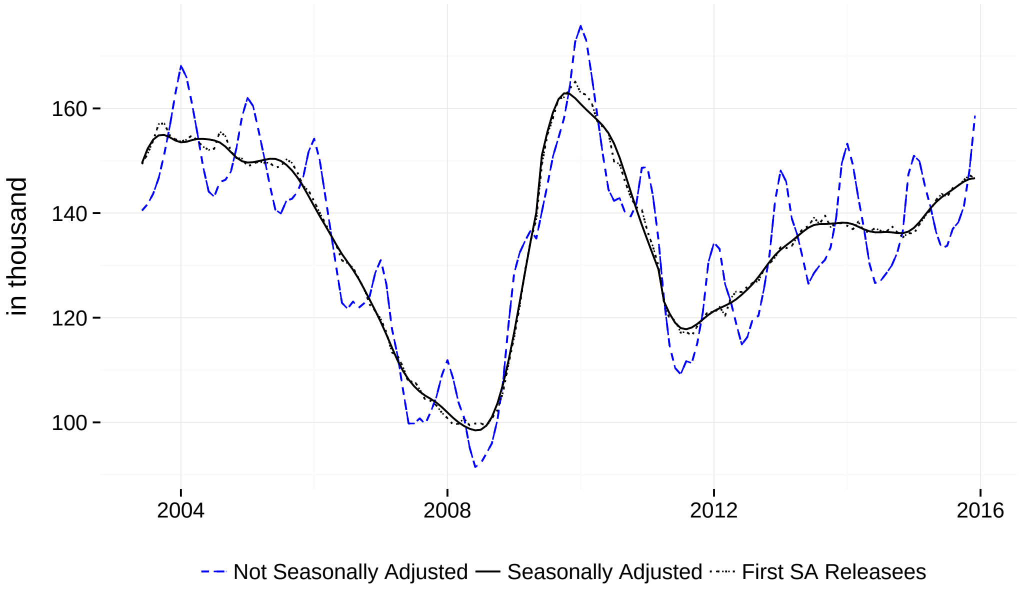


Figure 4.1 shows seasonally adjusted and unadjusted data for registered unemployed in Switzerland (latest SA vintage at the time of writing). In addition, the first releases of the seasonally adjusted figures are depicted. Whereas the unadjusted rawdata is never revised, the whole seasonally adjusted series experiences revisions whenever a new observation gets added. Compared to the level of the series, these revisions are almost negligible. The mean absolute revision to the seasonally adjusted level amounts to 784 persons, which corresponds to about 0.6% of the level on average. The maximum revision (3426 unemployed, occuring in 2010m2) corresponds to about 2.3% of the seasonally adjusted level in the same period. However, looking

full transparency about the seasonal adjustment methods and hence, it is possible to reconstruct longer series. Also note that the official "seasonally adjusted" figure consists only of the trend-component (equalling the seasonally adjusted series minus an irregular component). Naturally, a slightly smoother "seasonally adjusted" series results from this approach. This differs to the procedure in most other countries with the exception of the Australian Bureau of Statistics which publishes the trend-estimates *along* with the seasonally adjusted data for many variables. The results in this paper are based on the official "seasonally adjusted" figures but they carry to a large extent over to the actual seasonally adjusted series that includes the irregular component.

at unemployment growth, which is usually the focus of policy makers in the short term, revisions become strikingly more important. During 2003m6 and 2015m12, unemployment changed on average by 1'113 persons in absolute terms. The average absolute revision to that change amounted to 601 unemployed. That is, the revision to the preliminary release is on average about half as big as preliminary release itself. For 27% of the observations, the revision was bigger than the initially published change (the mean absolute revision for these cases is 868) and 18% of all initial announcements are indicating the wrong sign for the change of unemployment (the mean absolute revision for these cases is 895). Similar but somewhat less extreme results can be found for US employment (Table 4.1).

Table 4.1: Revision Summary Statistics for Swiss Unemployment and US Employment

	Swiss Unemployment	US Employment
Mean absolute change (first differences)	1'113	222'486
Mean abs. revision of the change	601	84'945
Maximum abs. revision of the change	3'190	335'000
Wrong preliminary sign	18%	9%

Note: Revisions are calculated with respect to the latest vintage available at the time of writing. All calculations are based on first differences (in persons) with respect to the previous month. In accordance with the analysis below, the most extreme preliminary releases were excluded (mostly they are from the extraordinary time period of the great recession)

A popular way to characterize such data revisions is to model them as *news* or *noise* as suggested in Mankiw, Runkle and Shapiro (1984) and Mankiw and Shapiro (1986). According to the news view, the preliminary release of an observation, x_t^p, is an efficient (rational, optimal) forecast of its final (true) release, x_t^f.[5] In the news case, the error term ν_t in Equation (4.1) reflects new information that is incorporated after the data has been released

[5] The literature has debated the question, when a data release can be considered final or true. For some economic time series, the revision process potentially continues indefinitely as there are many different reasons for revisions (see e.g. Indergand and Leist, 2014). Since benchmark revisions often change the definition of a variable, the last release of a variable before a benchmark revision is often considered as the final estimate and, assuming that all revisions bring a variable closer to its true state, can also be considered to be closest to the true value (see the discussion and references in Chapter 3).

for the first time:

$$x_t^f = x_t^p + \nu_t. \tag{4.1}$$

This requires the error term (i.e., the revision to the preliminary release, $rev_t^f = x_t^f - x_t^p$) to be orthogonal to all information available at release date p, most notably to the initial release x_t^p itself.

On the other hand, the error term (the revision) is said to contain noise if it is orthogonal to the final release x_t^f but correlated with the preliminary release x_t^p. In this case, the preliminary release is a suboptimal (inefficient, irrational) forecast of the final release and hence, the optimal estimate would involve a transformation of x_t^p using information that is known at p.

Hence if the noise hypothesis is true, we should find $\hat{\beta}_1 \neq 0$ and $\hat{\beta}_2 = 0$ in the following regressions:

$$rev_t^f = \alpha_1 + \beta_1 x_t^p + \epsilon_t, \tag{4.2}$$

$$rev_t^f = \alpha_2 + \beta_2 x_t^f + \delta_t. \tag{4.3}$$

On the other hand if the news hypothesis is true, we should find $\hat{\beta}_1 = 0$ and $\hat{\beta}_2 \neq 0$.[6]

Table 4.2 shows the results for these two regressions in the case of Swiss registered unemployed as well as US employment from the household survey. I use first differences rather than growth rates, as the former generally receive more attention by the public. The most extreme preliminary releases (more than two standard deviations away from the mean) were excluded from the regressions. Througout the paper, I use autocorrelation and heteroskedastic-

[6] Of course, there are intermediate cases possible where both hypotheses get rejected. In the original news-noise framework, there is no guidance for this case, however see Jacobs and Van Norden (2011) for a generalization. Also note that forecast rationality may also be violated if other variables known at the time of the data release can be used to predict future revisions. However, I consider this highly unlikely in the case of the two variables in question as there are no other sources for revisions other than the seasonal filters.

Table 4.2: Forecast Efficiency Test

	Swiss Reg. Unemployed			US Employment (Househ. Surv.)		
	Bias	Eq. (4.2)	Eq. (4.3)	Bias	Eq. (4.2)	Eq. (4.3)
Intercept	-43.96	-97.36	-27.03	1465.75	17696.45	2396.68
	(69.84)	(94.94)	(71.88)	(7275.68)	(9095.45)	(9255.94)
Prel. Release ($\hat{\beta}_1$)		-0.23^{**}			-0.16^{***}	
		(0.08)			(0.03)	
Final Release ($\hat{\beta}_2$)			0.06			-0.01
			(0.05)			(0.04)
Adj. R^2	0.00	0.18	0.00	0.00	0.16	-0.01
W($k = 2$)		8.30^{**}	0.40		24.10^{***}	0.07
Degr. of Freedom	141	140	140	145	144	144

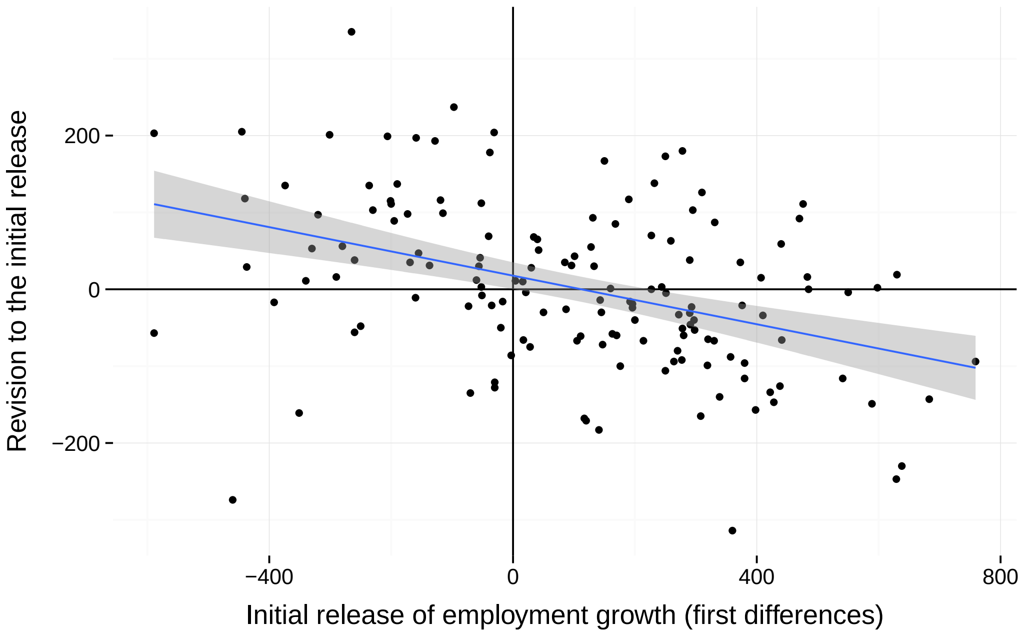
Note: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Dependent variable: Revision between first and latest available release. $N = 142$ for Swiss Reg. Unempl. and $N = 146$ for US Employment. Newey and West (1987) standard errors with 3 lags. The most extreme preliminary releases (larger than 2 standard deviations) were omitted from the regressions. W depicts the Wald test statistic for joint significance of the two coefficients. W is asymptotically χ^2_k distributed with $k = 2$ degrees of freedom.

ity robust standard errors (Newey and West, 1987) with appropriate lags.[7] Columns 1 and 4 show regressions on a constant, testing whether the revisions are unconditionally biased. Columns 2 and 5 show the results for Equation (4.2) and Columns 3 and 6 adress Equation (4.3).

Unsurprisingly, seasonal adjustment does not introduce a systematic bias into the adjusted data: The revisions have a zero mean. More surprising may be tha fact that the revisions are undoubtedly classified as mean-reverting noise rather than news in both cases. As illustrated in Figure 4.2, this means that high initial releases of the employment change tend to be negatively revised in later releases whereas low initial announcements are likely to be revised upward (see also the corresponding Figure 4.C.2 for Swiss registered unemployed).

[7] Lags are chosen based on the apparent autocorrelation and heteroskedasticity in residuals and alternative lag choices were always tested. In this case, only mild patterns could be detected and a lag of 3 was used. Interestingly, however, the results for Swiss registered unemployed change slightly when growth rates are used instead of first differences. In this case, both coefficients for the news and noise hypotheses are borderline significant on the 10% significance level. However, the adjusted R-squared in the news case remains essentially zero.

Figure 4.2: US Employment: Mean-Reversion of Initial Announcements



Note: Data in thousand employees. Revisions are calculated based on the latest available vintage. The line depicts the beta coefficient in Equation (4.2) (see Table 4.2). The most extreme positive and negative first releases were omitted.

The fact that first releases of SA data can sometimes be classified as irrational, but NSA releases cannot, is not completely new. Kavajecz and Collins (1995) find rationality for preliminary U.S. money stock growth rates in the case of NSA data but irrationality when SA data is used. They show that the X-11-ARIMA algorithm produces suboptimal preliminary SA releases. This inefficiency may be reduced by using concurrent seasonal adjustment, that is, continuously updating the seasonal components. However, the authors remain inconclusive on whether irrationality vanishes in the SA series if optimal seasonal adjustment is performed (concurrent adjustment with a correct specification of the DGP). Section 4 presents a simulation study to answer this question. In particular, X-11 type adjustment is compared to model-based adjustment (SEATS). The latter clearly delivers superior results both with regard to the size of revisions and forecast rationality. The following section presents some background on the two approaches.

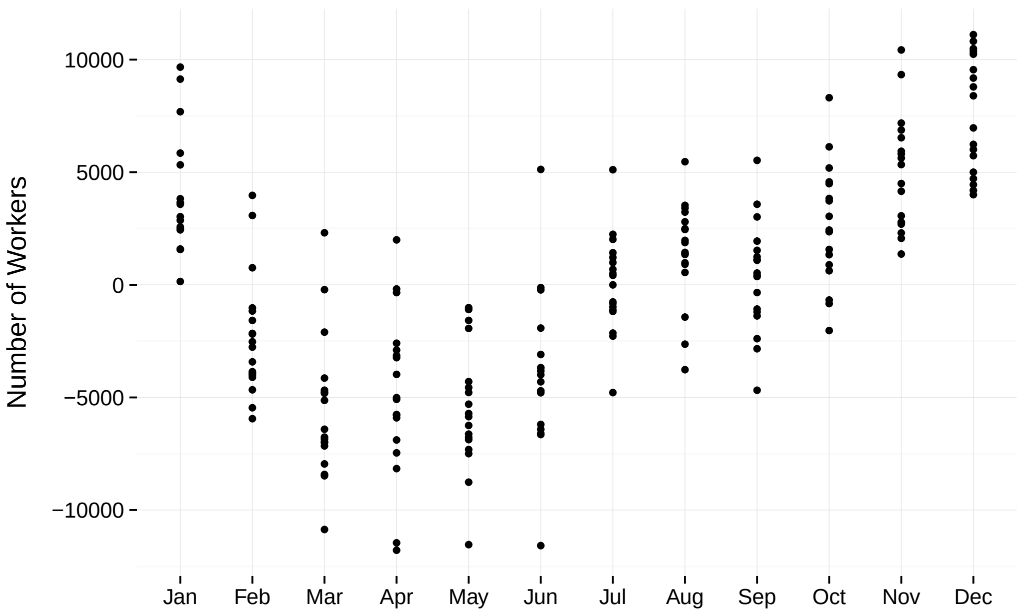
4.3 Background on Seasonal Adjustment

4.3.1 A Note on the Nature of Seasonality

Seasonal effects arise due to several reasons (which may not be completely distinct). The first is natural: Regular fluctuations in weather variables such as snowfall, rainfall or temperature can boost or hamper economic activity. Some industries are much more affected by weather fluctuations than others[8] and the effects may be opposing for different sectors. Farming, construction and tourism are typically positively affected by nice and warm weather whereas the energy or health care sector are positively affected by harsh winters. The second reason is institutional: The timing of school holidays and university semesters influences travel behaviour of large parts of the population as well as job market entries of graduates. The end of a tax year or an accounting period (influencing dividend and bonus payments) on the other hand affect disposable incomes. The third reason is cultural: Consumer behaviour is heavily influenced by holidays, most notably Christmas for Western countries. As a result, retail sales surge towards the end of the year and fall at the beginning of the year. For the US, Barsky and Miron (1989) find that Christmas is generally the most important cause for seasonality in economic time series. Again however, its impact may be very distinct for different variables. For example, goods consumption in the US displays a heavy spike in the fourth quarter whereas service consumption hardly displays any seasonal pattern (Miron, 1996). In addition to these three basic causes, the expectation of such seasonal patterns can cause actual seasonality in other variables (such as prices, inventories, wages, employment) since production and consumption plans are adjusted accordingly (Granger, 1979). Often, seasonality in a variable may be caused by a combination of all these reasons or by indirect impacts through other variables. For example, local consumption may be affected by the number of guest workers which may

[8] Boldin and Wright (2015) estimate the effect of unseasonal (extraordinary) weather on employment data. They find that deviations of the temperature and snowfall from seasonal norms affect sectors very differently with the highest impact in the construction sector.

Figure 4.1: Swiss Unemployment, First Differences (1998m2 - 2015m12)



be affected by construction plans and tourism activities which in turn are affected by weather conditions.

The interplay of all these sources may be complex, however, the arising seasonal fluctuations often seem to be relatively regular. For example, Figure 4.1 shows the growth of not seasonally adjusted Swiss registered unemployed in first differences seperately for each month. Unemployment exclusively grows from November through January and almost exclusively falls from March through June during the whole time period.

Therefore it might be reasonable to assume seasonal patterns to be stationary and deterministic. The conventional representation of such a process is (see for example Miron, 1996 or Ghysels and Osborn, 2001):

$$x_t = \sum_{s=1}^{S} \alpha^s d_t^s + \epsilon_t \tag{4.4}$$

where d_t^s represents a dummy for each season s with S being the total number of seasons (4 in the quarterly case and 12 in the monthly case).

α^s is the mean of variable x_t in season s and ϵ_t is a stationary stochastic process with mean zero. The variable x_t is typically a detrended version of the (nonstationary) level X_t. The upper panel in Table 4.1 shows the results from applying Equation (4.4) on GDP data from nine countries.[9] Confirming the results in Beaulieu and Miron (1992), seasonal fluctuations are very important, explaining between 0.7 and 0.9 of the variability of GDP in most countries. Furthermore, the seasonality appears to be relatively similar across borders. In most countries, there is a first quarter slump after the rise of output around Christmas. In addition, Italy and France seem to display a decrease of the unadjusted GDP in the third quarter which is most likely related to summer vacation of the local population. Interestingly, both Germany and Switzerland display different seasonal patterns with no strong spike in the fourth quarter. However, splitting the sample in two in 1991 reveals that the seasonal pattern changed over time for these two and to a lesser extend also for other countries, see the lower two panels in Table 4.1 (when splitting the data in 1991, Italy and Japan fall out of the sample as no unadjusted GDP data exists for these two countries prior to 1995).[10] Furhermore, with the exception of Norway, the adjusted R^2 rises substantially for most countries when considering the periods seperately. It may therefore be important to model the series in a way as to allow for changing seasonals.

This is what the second conventional representation for seasonal effects accomplishes: The non-deterministic (stationary or nonstationary) representation of seasonal effects is a generalization of autoregressive moving-average (ARMA) models to the seasonal frequency. The simplest example of such a

[9] I obtained the data for these countries from the OECD database. For most other countries, not seasonally adjusted GDP figures are available for a much shorter time period and therefore were not considered.

[10] 1991 is the date of the German unification. In Switzerland, GDP data before 1990 gets retropolated using older figures that were obtained under a somewhat different National Accounts measurement system (Indergand and Leist, 2014). It might therefore not be very surprising if the seasonals may have changed around 1990.

Table 4.1: Seasonal Patterns in GDP

	Q1	Q2	Q3	Q4	adj. R^2	N
			Full Sample			
UK	-2.93	-1.40	3.44	3.09	0.78	163
JPN	-3.61	-1.76	2.33	3.83	0.76	87
ITA	-5.90	4.42	-1.88	3.69	0.91	79
GER	-2.99	1.69	2.26	1.11	0.54	183
FRA	-2.21	1.17	-2.92	5.68	0.91	143
CHE	-1.46	1.32	0.85	1.01	0.34	143
KOR	-13.37	10.95	0.68	8.64	0.76	183
AUS	-10.60	3.46	1.69	8.80	0.88	223
NOR	-2.26	-1.73	-0.40	7.00	0.78	151
			Data until 1991			
UK	-4.32	-0.99	3.88	3.94	0.84	63
GER	-4.49	2.11	1.96	3.27	0.78	83
FRA	-2.86	1.50	-4.06	7.70	0.98	43
CHE	1.15	1.13	1.11	-1.25	0.65	43
KOR	-19.36	15.02	0.54	12.83	0.86	83
AUS	-13.09	3.26	1.99	11.40	0.95	123
NOR	-1.92	-1.39	-0.37	6.82	0.73	51
			Data after 1991			
UK	-1.92	-1.66	3.15	2.56	0.80	99
GER	-1.86	1.33	2.51	-0.71	0.60	99
FRA	-1.90	1.03	-2.42	4.80	0.93	99
CHE	-2.61	1.40	0.73	2.00	0.69	99
KOR	-8.57	7.53	0.79	5.12	0.95	99
AUS	-7.44	3.71	1.33	5.58	0.96	99
NOR	-2.40	-1.91	-0.42	7.10	0.79	99

Note: OLS-Results for Equation (4.4): Regression on on seasonal dummies. Growth rates in %.

process is the first-order seasonal autoregressive process:

$$x_t = \Phi^s x_{t-s} + \epsilon_t. \tag{4.5}$$

As above, ϵ_t is a stationary disturbance with a zero mean. Analoguous to common $AR(1)$ processes, this process can be represented as the sum of all innovations by recursive substitution. Seasonality can then be seen to originate from the unobserved starting value for season s and from the stochastic shocks that tend to be repeated but diminish over time if $|\Phi^s| < 1$. On the other hand, if $|\Phi^s| = 1$ (the seasonal random walk), then the seasonal means are not defined. Under such a process, random shocks do not have a tendency to return to the mean and, as Miron (1996) notes, the seasonal effect of Christmas would be allowed to shift into the summer. However, the presence of seasonal unit roots would lead to spurious coefficients in Equation (4.4). It is thus very important to know whether the seasonal process should be characterized as stochastic and wether seasonal patterns change over time.

There are many reasons why seasonal patterns in case of all three mentionned causes may change over time. For example, climate change may result in warmer weather on average und thus the seasonal impact of the winter may diminish. Shifts in preferences such as secularization may diminish the importance of some holidays in favor of others. Increasing trade links with Asia may increase the importance of foreign holidays such as Chinese New Year or Indian Diwali for trade variables with according effects on production of export-oriented sectors. The seasonal patterns in tourism-related variables will also change in Western countries with booming visits of Asian guests. All this is likely to introduce considerable instability into the seasonal patterns of many variables.

Several tests for seasonal stability have been proposed in the literature. Dickey, Hasza and Fuller (1984) provide a extension of Dickey and Fuller (1979) unit root test to the seasonal case. It tests the null of a seasonal unit root, $\Phi = 1$, in

$$(1 - \Phi L^s)x_t = \epsilon_t, \tag{4.6}$$

with stationary ϵ_t, and uses the auxiliary regression

$$(1 - L^s)x_t = \gamma x_{t-s} + \epsilon_t. \tag{4.7}$$

The null of seasonal integration corresponds to $\gamma = 0$. Simulated critical values for this test are provided by Dickey, Hasza and Fuller (1984). However, this does not allow to test for unit roots at some but not all seasonal frequencies. Hylleberg et al. (1990) provide a generalization (henceforth HYGE test) that overcomes this disadvantage. For the quarterly[11] case, their test is based on the auxiliary regression

$$(1 - L^4)x_t = \rho_1 x_{t-1}^{(1)} + \rho_2 x_{t-1}^{(2)} + \rho_3 x_{t-2}^{(3)} + \rho_4 x_{t-1}^{(4)} + \epsilon_t, \tag{4.8}$$

where $x_t^{(1)} = (1 + L + L^2 + L^3)x_t$ removes seasonal unit roots (but leaves unit roots at the zero frequency), $x_t^{(2)} = -(1 - L + L^2 - L^3)x_t$ removes unit roots at frequencies other than π, $x_t^{(3)} = -(1 - L^2)x_t$ leaves roots at the other seasonal frequencies, $\pi/2$ and $3\pi/2$. This auxiliary regression can be estimated by OLS and coefficients on the ρ's can then be tested to exhamine the nature of the seasonal effects. The t-statistics on the ρ's are distributed as noted in Hylleberg et al. (1990). More specifically, $\rho_2 = 0$ tests for a seasonal unit root at the frequency π whereas unit roots at frequencies $\pi/2$ and $3\pi/2$ imply $\rho_3 = \rho_4 = 0$ jointly. Table 4.2 shows the results for testing the null hypothesis of $\rho_2 = 0$ in Column 1, $\rho_3 = \rho_4 = 0$ in Column 2 and $\rho_2 = \rho_3 = \rho_4 = 0$ (no seasonal unit root) in Column 3. I test GDP data for all nine countries as well as data for Swiss registered unemployed and US employment from the household survey. As the latter two are of monthly frequency, the table only shows results for the joint test of all seasonal frequencies. The evidence is mixed. Whereas the null of a seasonal unit root in the case of the biannual frequency (π) gets not rejected in most cases, the joint tests generally reject the null and point to stationarity. In the case of Norway, Swiss registered unemployed as well as US employment, seasonal unit roots get overwhelmingly rejected.

Hence in most cases the null of a seasonal unit root, $\Phi = 1$ in Equation (4.5), is rejected by the HEGY test. However, the test has relatively low power in small samples as Canova and Hansen (1995) note. Therefore they suggest to test null of deterministic seasonality, generalizing the test for stationarity by Kwiatkowski et al. (1992) to the seasonal case (henceforth CH test). They develop a Lagrange multiplier test statistic for the null hypothesis of no unit roots at specific (or all) seasonal frequencies. Furthermore, their test also allows to test wether individual (or all) deterministic inter-

[11] Beaulieu and Miron (1993) generalize this test to monthly data.

Table 4.2: HEGY Test for Seasonal Unit Roots

	ρ_2	$\rho_2 \cap \rho_3$	$\rho_2 \cap \rho_3 \cap \rho_4$
UK	-2.99*	7.55*	7.87***
GER	-2.24	13.93***	11.02***
FRA	-1.68	9.23**	7.16***
CHE	-1.36	8.21**	6.15***
KOR	-1.73	7.79**	6.21***
AUS	-2.53	7.26*	6.89***
NOR	-6.04***	52.21***	45.36***
CHE Unemployment			62.23***
US Employment			35.71***

Note: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Col. 1 tests seasonal unit roots (null hypothesis) at frequency π. Col. 2 tests at frequencies $\pi/2$ and $3\pi/2$. Col. 4 jointly tests on all frequencies. GDP data if not labelled otherwise. The auxiliary regression includes seasonal dummies.

cepts are stable against the alternative hypothesis of non-constant seasonal intercepts.

Again, I test all real-world series considered in this paper, that is, quarterly GDP data for nine countries as well as Swiss registered unemployed and US emplyoment from the household survey (since the latter two are on a monthly frequency, I report only the joint test for breavity).[12] Table 4.3, shows the results for the CH test. Following Canova and Hansen (1995) and others, I include one lag of the dependend variable in the tests in order to reduce the serial correlation in the residuals (including more lags may absorb some of the seasonal unit roots). The first four columns show the results for testing the null of deterministic seasonality for each quarter seperately, Column 5 shows their joint significance. With the remarkable exception of Norway, all countries show changing seasonal patterns in the case of two or more quarters. Column 6 and 7 test the null of no unit root at the seasonal frequencies and a joint-test is displayed in Column 8. For all countries with the exception of Norway, the tests clearly point to seasonal unit roots.

Overall, the evidence is mixed on whether (stationary) stochastic seasonality is present or not. For GDP data of Norway, the tests favor deterministic

[12] In case of I(1) series the CH-test is applied to the log first differences. I use the standard augmented Dickey-Fuller and KPSS tests to test for stationarity.

Table 4.3: Canova-Hansen Test for Seasonal Stability and Seasonal Unit Roots

	Q1	Q2	Q3	Q4	Joint	π	$\pi/2$	Joint
UK	1.27***	0.24	1.07***	1.29***	2.37**	2.02***	0.93**	2.27***
GER	0.83**	0.14	0.3	2.19***	2.94**	2.65***	2.12***	2.86***
FRA	0.31	0.12	1.87***	2.15***	2.85**	2.37***	2.34***	2.76***
CHE	1.16***	0.63*	0.05	1.55***	2.09**	1.79***	1.8***	2.01***
KOR	1.54***	0.09	1.92***	2.28***	3.01**	2.53***	1.65***	2.74***
AUS	1.96***	2.4***	0.74**	3.04***	3.78**	3.45***	2.06***	3.67***
NOR	0.2	0.18	0.17	0.08	0.72	0.19	0.11	0.37
CHE UE					2.02			1.93
US EMP					4.73*			4.7***

Note: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Columns 1-4 show the test statistics for each quarter, Column 5 the joint test. Columns 6-7 show the test statistics for the seasonal frequencies, Column 8 the joint test. GDP data if not labelled otherwise

seasonality as well as for Swiss unemployed. In other cases, the evidence is mixed. The HEGY test often rejects the presence of a seasonal unit root whereas the CH test rejects stationarity, which of course can happen in small samples. This confirms the results of the literature that has remained somewhat inconclusive.[13] Note, however, that the HEGY test is for the null of a seasonal unit root. Hence, rejection does not necessarily favor deterministic over stationary stochastic models, which the CH test explicitly accounts for. Furthermore, stochastic seasonality remains the overwhelmingly most common assumption on which statistical agencies around the globe rely: Explicitly in the case of model-based adjustment (with a generalized version of Equation (4.5), see Section 4.32) and implicitly in the case of moving average based adjustment (where the length of the seasonal moving-average filter determines the degree of change allowed in the seasonals). In the remainder of this paper, I will focus on these adjustment approaches and on series that fulfill this assumption.

[13] For example, Beaulieu and Miron (1993) and Osborn (1990), relying mainly on the HEGY test, conclude that the majority of US and UK time series can be described as deterministic whereas some slightly more recent studies such as Canova and Hansen (1995) and Canova and Ghysels (1994) find nonstationary seasonal effects in most cases.

4.3.2 Methods and Programs for Seasonal Adjustment

Generally, there are two approaches for seasonal adjustment, the non-parametric and the parametric. The first estimates the seasonal components essentially by taking moving averages of the data in the same season over different years, after the trend has been removed from the original series. This class of seasonal filters was introduced by the U.S. Census Bureau in 1957 and further developments lead to the publication of the X-11 algorithm (Shiskin, Young and Musgrave, 1965). The algorithm has been widely used and further developed, giving rise to the X-11-ARIMA program (Dagum, 1980) and enhanced versions thereof (such as X-11-ARIMA/88, see Dagum, 1988), and the X-12-ARIMA program (Findley et al., 1998). All these further developments of the X-11 *algorithm* are often called the X-11 *family*. Various other non-parametric techniques have been developed in the past (see the references in Burman, 1980). However, most of them are not anymore in frequent use by major statistical agencies, a notable exception is the German BV4.1 procedure.

The second, somewhat younger type of seasonal filters relies on a parametric, model-based approach. In this procedure, the underlying DGP is explicitly modelled and the components are inferred using signal extraction techniques by either using a state space representation or relying on frequency domain analysis. While the theoretical foundations have been developed earlier, Burman (1980) was among the first to provide an operational procedure that was applied by the Bank of England. In 1989, the new STAMP (Koopman et al., 1996) program incorporated model-based adjustment but was applied only by few official statistical agencies. Based on the Burman program, TRAMO-SEATS was developed at the Bank of Spain (Gomez and Maravall, 1996) and is the most prominent implementation of model-based adjustment today. It got widely applied since 1994 by Eurostat and many small European countries as well as Italy and Spain.

The program Demetra+ published by Eurostat in 2007 provides a unified interface for both X-12-ARIMA and TRAMO-SEATS. More recently, the two seasonal adjustment approaches have culminated into the X-13ARIMA-

SEATS program developed by the U.S. Census Bureau in collaboration with the Bank of Spain. The software is freely available on the Census Bureau's Website and allows the user to choose between a parametric (SEATS) and a non-parametric (X-11) decomposition. The implementation of the X-11 decomposition in X-13ARIMA-SEATS is essentially the same as in the X-12-ARIMA program.

The X-11 approach probably remains as the most applied procedure to date. In 2002, about 90% of OECD countries used a method of the X-11 family though that share was expected to decline in favor of TRAMO-SEATS (OECD, 2002).

In what follows, I provide a short overview about these approaches. As both are rather involved, I keep this discussion brief. For a detailed description of X-12-ARIMA (and the X-11 algorithm), see the introductory paper by Findley et al. (1998). For details about the SEATS decomposition see Gomez and Maravall (1996). Ghysels and Osborn (2001) provide a broader overview about the literature and the different approaches.

Seasonal adjustment means extracting the seasonal regularities, y_t^s, from a time series y_t. To do so one assumes that the series y_t can be decomposed multiplicatively into a trend, y_t^t, a seasonal factor, y_t^s, and an irregular factor, y_t^i, as

$$y_t = y_t^t \times y_t^s \times y_t^i. \tag{4.9}$$

The seasonally adjusted series then equals the unadjusted series divided by the seasonal factor,

$$y_t^{sa} = \frac{y_t}{y_t^s} = y_t^t \times y_t^i. \tag{4.10}$$

The multiplicative formulation in Equation (4.9) is appropriate for many macroeconomic time series as the seasonal variations often increase proportionally with the level of the series. For the additive formulation that assumes constant seasonal variations over time, see Ghysels and Osborn (2001).

Before the decomposition, y_t is usually purged from outlier, working day and holiday effects. These estimated effects are then added back to the series after seasonal adjustment (however, statistical agencies often publish seasonally and working day adjusted series, also excluding the working day effect). In the X-13ARIMA-SEATS program, very similar to its predecessor softwares X-12-ARIMA and TRAMO-SEATS, the estimation of these effects is done in a preadjustment step by fitting a linear model to the raw series, x_t,

$$x_t = \sum_{i=1}^{n} \beta_i z_{i,t} + y_t. \tag{4.11}$$

$z_{i,t}$ may contain dummy vectors for additive outliers, level shifts, or specific holidays (such as Easter or Chinese New Year), a vector indicating the number of working days or other user specified regressors, for example in order to estimate weather effects, see Boldin and Wright (2015). y_t is assumed to follow a seasonal autoregressive integrated moving-average process, $SARIMA(p,d,q)(P,D,Q)^s$, written as

$$\phi_p(L)\Phi_P(L^s)(1-L)(1-L^s)y_t = \theta_q(L)\Theta_Q(L^s)\xi_t, \tag{4.12}$$

where y_t is given in (4.11) and $\phi_p(L)$ and $\Phi_P(L)$ are the non-seasonal and the seasonal autoregressive polynomials of order p and P, respectively. $\theta_q(L)$ and $\Theta_Q(L)$ stand for the non-seasonal and the seasonal moving-average polynomial of order q and Q. $(1-L)$ takes first and $(1-L^s)$ seasonal differences with s representing the seasonal period where L is the lag operator, $Ly_t = y_{t-1}$. Finally, ξ_t is assumed to be white noise.

The model in Equation (4.12) that was fitted to y_t is used to forecast and backcast y_t in order to improve the seasonal adjustment at the beginning and the end of the series. This is important since symmetric instead of asymmetric filters should be used in order to reduce the size of revisions of the resulting series (Dagum and Laniel, 1987). X-13ARIMA-SEATS uses the same peradjustment technique for both model-based and moving-average adjustment and then hands the preadjusted, for- and backcasted series over to SEATS or X-11 for decomposition.

For the actual seasonal adjustment steps X-11 discards the estimated model in Equation (4.12) and instead sequentially applies various moving average filters. To start, an initial decomposition is performed where the series is detrended using a centered 13-term moving average (Step 1). Then, a 3x3 seasonal moving average and another centered 13-term moving average are applied to the detrended series to get the initial estimate of the seasonally adjusted series (Step 2). Next, a $2H + 1$ term Henderson filter (with data-determined H) is applied to the initial seasonally adjusted series to get the intermediate detrended series (Step 3), which in turn is processed once more through a centered 13-term moving average and a 3x5 seasonal moving average in order to arrive at the "final" seasonally adjusted series (Step 4). From this, the "final" trend and irregular series can be calculated (Step 5). Then, an outlier-correction is applied to the irregular series (Step 6). With the original data replaced by the product of the "final" trend, the "final" seasonal and the outlier-corrected irregular factors, all six steps are repeated two more times (with slightly adjusted Steps 3 and 4, see Appendix 1 for more details and references).

Hence, eventough X-11 allows for some data-driven modifications, it ignores the specific $SARIMA(p, d, q)(P, D, Q)^s$ structure underlying the DGP of the series, as specified in Equation (4.12) and fitted to the raw data prior to the decomposition.

On the other hand, model-based adjustment views the seasonal decomposition as a signal extraction problem where the $SARIMA$ model fitted to y_t is explicitly used for the decomposition. From the estimated model for the DGP of y_t, appropriate ARIMA processes for y_t^s and y_t^{ns} can be derived if identifying restrictions are imposed (these restrictions may not be harmless, see Ghysels and Osborn, 2001). This yields a fully data-driven modelling of the factors. In this tradition, two approaches have emerged to extract the unobserved time series. The first formulates the problem in state-space form and uses the Kalman filter. This approach is described for example in Ghysels and Osborn (2001) and implemented in the STAMP program. The second approach relies on the Wiener-Kolmogorov filter which provides the minimum mean square error estimator for each component. This approach is implemented in the SEATS program and for example described in Gomez and Maravall (2001).

Contrary to the X-11 approach, model-based adjustment allows for a decomposition explicitly based on the underlying (estimated) DGP of the NSA series. Intuitively, it seems therefore natural that the model-based estimates of the SA series converge more quickly to their final values, provided that the model is correctly specified. As noted in the conclusion of Burridge and Wallis (1984), the differences between X-11 and signal extraction techniques will depend on the order and the coefficients of the polynomials in the DGP of the series, characterized by Equation (4.12). Moreover, the traditional X-11 filter happens to be optimal for a particular type of model in a mean-squared error signal extraction sense (see Cleveland, 1972; Burridge and Wallis, 1984 and Maravall, 1985). For any other DGP, X-11 will be suboptimal. Burridge and Wallis (1984) show for symmetric and asymmetric filters that this model consists of as many as 26 moving average terms with at least 13 non-negligible coefficients. Depoutot and Planas (2002) show that a much more parsimonious model considered in the simulation excercise below also delivers optimal filters that are very close to the X-11 filters embodied in the X-12-ARIMA program.[14]

Intuitively, revisions to the preliminary estimates may hence be similar for the particular model for which X-11 and signal extraction techniques are (almost) equivalent and less similar for other models. To the best of my knowlege, however, no study exists that compares these differences from an applied perspective, particularly with respect to the forecast efficiency of preliminary releases. With respect to the latter, there is little guidance on whether these methods, as implemented in the X-13ARIMA-SEATS program, perform differently. The next section sheds light on this issue.

[14] As described above, these X-11 variants allow the user to choose between a 3x3, 3x5, 3x9 and 3x15 seasonal moving average as well as between different Henderson filters. For all these cases, Depoutot and Planas (2002) provide coefficients for which X-11 and SEATS are very close.

4.4 Results from a Simulation Excercise

In order to investigate the effects of different adjustment methods on the resulting data, I simulate a large number of $SARIMA(P, D, Q)(p, d, q)$ series. I then perform seasonal adjustment in real-time using both a moving-average based (X-11) and a model-based (SEATS) filter. I then calculate revisions to preliminary SA releases in order to test for forecast rationality and to compare the size of revisions. For all simulations, I use DGPs of the form

$$(1 - L)(1 - L^s)y_t = (1 + \theta)(1 + \Theta L^s)\xi_t, \tag{4.13}$$

where θ and Θ are MA parameters with modulus smaller than unity. This is the famous airline model which fits the majority of macroeconomic time series well (Fischer and Planas, 2000) and may be the most widely applied model for seasonal adjustment by statistical agencies. I focus on all combinations of the parameter values $\theta \in \{-0.7, -0.2, 0.2, 0.7\}$ and $\Theta \in \{-0.7, -0.2\}$.[15] While the choice of these parameters parameters is arbitrary, they include the range of parameters that is found when an Airline model is fitted to real-world data (I used GDP, employment and monetary aggregates and always found parameters between -0.2 and 0.7 for both θ and Θ). For each parameter pair, I simulate 500 series and seasonally adjust them using the X-11 and SEATS algorithms implemented in the X-13ARIMA-SEATS software provided by the Census Bureau.[16] Seasonal adjustment is done in a real-time setting for each of the 500 series, that is, first only the series for $t = 1..140$ is processed through the filter. This yields an initial seasonally adjusted estimate for $t = 140$, $\Delta sa_{t=140}^{v=1}$. Then one observation is added and the series $\Delta nsa_{t=1..141}^{v=2}$ is seasonally adjusted, yielding a first estimate for $\Delta sa_{t=141}^{v=2}$ and a revised estimate, $\Delta sa_{t=140}^{v=2}$ for $\Delta sa_{t=140}^{v=1}$. This procedure is applied to

[15] Since the airline model does not provide admissible decompositions for large positive Θ (Hillmer and Tiao, 1982), I provide only results for $\Theta < 0$. However, I also did simulations for $\Theta > 0$. In all cases, X-13ARIMA-SEATS was able to decompose the series. The results for these cases all point to the same direction as the ones presented here.

[16] `https://www.census.gov/srd/www/x13as/`. Sax (2015) provides a very helpful interface to the software.

each vintage, $v = 1..360$, yielding 360 preliminary estimates. I then calculate revisions with respect to the second release, $v+1$ (one revision), with respect to the 5th release, $v+5$ (4 revisions), with respect to the 30st release, $v+30$ (29 revisions), and with respect to the latest vintage, $V = 360$ (number of revisions varies for each observation). While for the latter, the number of revisions varies for each preliminary estimate, it corresponds most closely to the situation that a statistical agency or a policy maker faces in reality. As the DGP of the series are known, the seasonal filters can be optimally specified, both for the ARIMA forecast in the X-11 procedure and the SEATS decomposition. I therefore deactivated the automatical model detection as well as the automatical outlier and calendar days detection in the preadjustment step, Equation (4.11).[17] Furthermore, I use the 3x5 seasonal moving average for X-11 (but present alternative choices in the Appendix). Other than that, the default specification of the X-13ARIMA-SEATS program were used. This procedure corresponds to the practice of many statistical agencies which often fix the model, the outliers, calendar- and holiday adjustment as well as the seasonal moving average used in the X-11 algorithm. Usually, these choices get periodically reviewed (often once a year) and sometimes adjusted if the seasonal patterns have changed.

Table 4.1 shows revision statistics comparing the preliminary release to the 2nd, the 5th and the 30iest release, for the different DGPs and for both the X-11 and the SEATS procedure. For all DGPs, the mean average and mean squared revisions are unambiguously lower if the SEATS decomposition is used (Tables 4.C.1 and 4.C.2 in Appendix 3 show the results if a 3x3 and a 3x5 filter are used for the X-11 decomposition. I also did simulations for the 3x15 filter which confirm these results, but I omit these for brevity). This indicates that the preliminary estimates are generally closer to the final release, that is, more precisely estimated with model-based adjustment - at least for the default specifications of the filters that are often used both by statistical agencies and other data users. In some cases, the MAR is many-

[17] Note that X-11 includes an additional extreme value correction mechanism (see Appendix 1). For robustness, I reestimated all results by deactivating this mechanism, that is, immensely increasing the limits for an observation to be regarded as an outlier. The results entirely confirm the findings presented in this section.

Table 4.1: Average Mean Revisions

	Average Mean Absolute Revision					
	2. Release		5. Release		30. Release	
	X-11	SEATS	X-11	SEATS	X-11	SEATS
$\{\theta, \Theta\} = \{0.7, -0.2\}$	0.170	0.076	0.218	0.149	0.451	0.263
$\{\theta, \Theta\} = \{0.2, -0.2\}$	0.132	0.039	0.163	0.099	0.360	0.222
$\{\theta, \Theta\} = \{-0.2, -0.2\}$	0.110	0.021	0.132	0.062	0.346	0.231
$\{\theta, \Theta\} = \{-0.7, -0.2\}$	0.079	0.016	0.103	0.030	0.407	0.244
$\{\theta, \Theta\} = \{0.7, -0.7\}$	0.103	0.032	0.129	0.064	0.319	0.246
$\{\theta, \Theta\} = \{0.2, -0.7\}$	0.082	0.021	0.101	0.045	0.265	0.210
$\{\theta, \Theta\} = \{-0.2, -0.7\}$	0.068	0.014	0.085	0.030	0.259	0.215
$\{\theta, \Theta\} = \{-0.7, -0.7\}$	0.051	0.008	0.069	0.016	0.303	0.257
	Average Mean Squared Revision					
	2. Release		5. Release		30. Release	
	X-11	SEATS	X-11	SEATS	X-11	SEATS
$\{\theta, \Theta\} = \{0.7, -0.2\}$	0.083	0.010	0.118	0.036	0.359	0.111
$\{\theta, \Theta\} = \{0.2, -0.2\}$	0.045	0.003	0.063	0.016	0.225	0.079
$\{\theta, \Theta\} = \{-0.2, -0.2\}$	0.031	0.001	0.041	0.007	0.209	0.086
$\{\theta, \Theta\} = \{-0.7, -0.2\}$	0.019	0.001	0.028	0.003	0.291	0.095
$\{\theta, \Theta\} = \{0.7, -0.7\}$	0.023	0.002	0.031	0.007	0.167	0.096
$\{\theta, \Theta\} = \{0.2, -0.7\}$	0.016	0.001	0.020	0.004	0.115	0.070
$\{\theta, \Theta\} = \{-0.2, -0.7\}$	0.012	0.000	0.015	0.002	0.112	0.073
$\{\theta, \Theta\} = \{-0.7, -0.7\}$	0.008	0.000	0.012	0.001	0.155	0.105

Note: The table shows mean absolute revisions and mean squared revisions in percentage points, averaged over 500 simulated series.

fold larger for X-11 adjustment compared to SEATS. Furthermore, average revisions are in all cases much higher after 30 releases than after only two or five. This suggest that it takes several releases - in this case months - until the estimates converge close to their final value. For the SEATS decompositions, the revision statistics are much lower in the beginning but then increase more than in the case of X-11. This makes sense considering the fact that X-11 uses a relatively short filter (mostly a 3x5 seasonal moving average gets chosen), implying that revisions should become negligible after a 3 years, whereas seasonal adjustment with SEATS is in principle based on infinite (Wiener-Komogorov) filters.

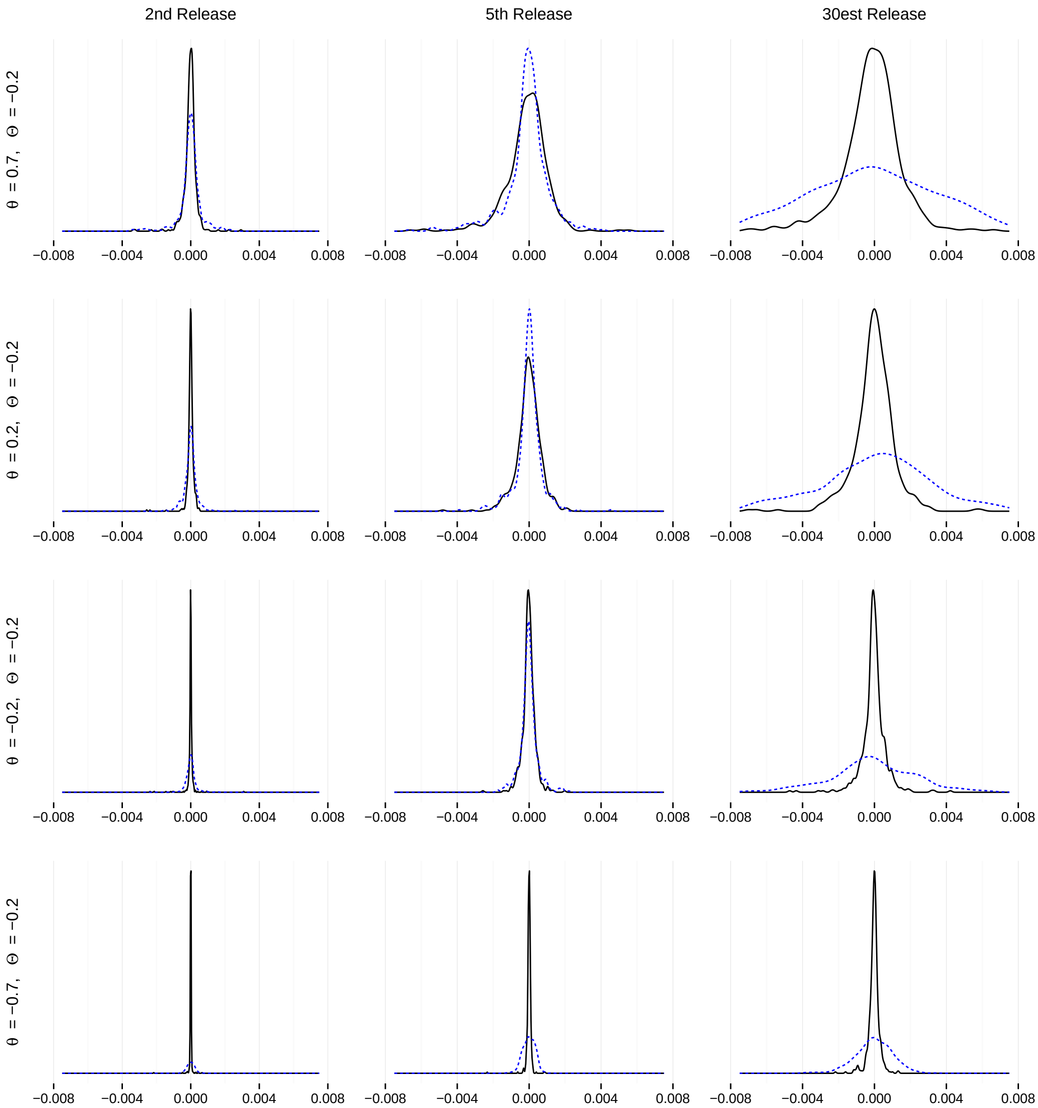
In what follows, I report the parameter distribution for $\hat{\alpha}$ and $\hat{\beta}$ in the forecast efficiency regression (Equation 4.2) for the different revision types, for the different DGPs and both for the X-11 and SEATS procedure. The densities are calculated using a Gaussian kernel density estimator. The bandwidth choice follows Silverman (1986) but the results are robust against other choices. Figures 4.1 and 4.2 show the density estimates for $\hat{\alpha}$. Both the X-11 (dotted line) and SEATS (straight line) based decomposition perform relatively well. That is, $\hat{\alpha}$ has in all cases a mean very close to zero. This corresponds to what Kavajecz and Collins (1995) have found for the X-11-ARIMA procedure that was used at the Federal Reserve to seasonally adjust the monetary base. However, model-based adjustment seems to reduce the variability of the estimate substantially, in-line with the results in Table **??**. For all DGPs and all revision types, the distribution of $\hat{\alpha}$ is centered closer around zero compared to the X-11 adjustment.

Yet a much more interesting picture emerges in case of the estimated densities for $\hat{\beta}$ displayed in Figures 4.3 and 4.4. The X-11 filter (dotted line) and to some extend also the SEATS filter (straight line) give rise to preliminary estimates that are too extreme. In later releases, they get revised towards their mean, that is, very large growth rates tend to be revised down and very low growth rates tend to be revised up, resulting in a negative $\hat{\beta}$ in Equation (4.2). This is the case for almost all specifications of the DGP that are adjusted with the default X-11 algorithm. For SEATS on the other hand, this tendency is by far less pronounced: The distributions of the β coefficients are only mildly skewed or shifted towards the left.

Moreover, it takes a considerable amount of time for these revisions to materialize. The variances of both $\hat{\alpha}$ and $\hat{\beta}$ increase a lot between the 5th and the 30iest release for both SEATS and X-11. Mean reversion of preliminary estimates accordingly does not take place until many releases have past. In fact, the initial estimas for β in the case of X-11 adjustment have a positive mean in all cases (indicating the opposite tendency, that is, positive releases tend to be corrected upward and vice versa) but then shift into the negative area after more than 4 revisions.

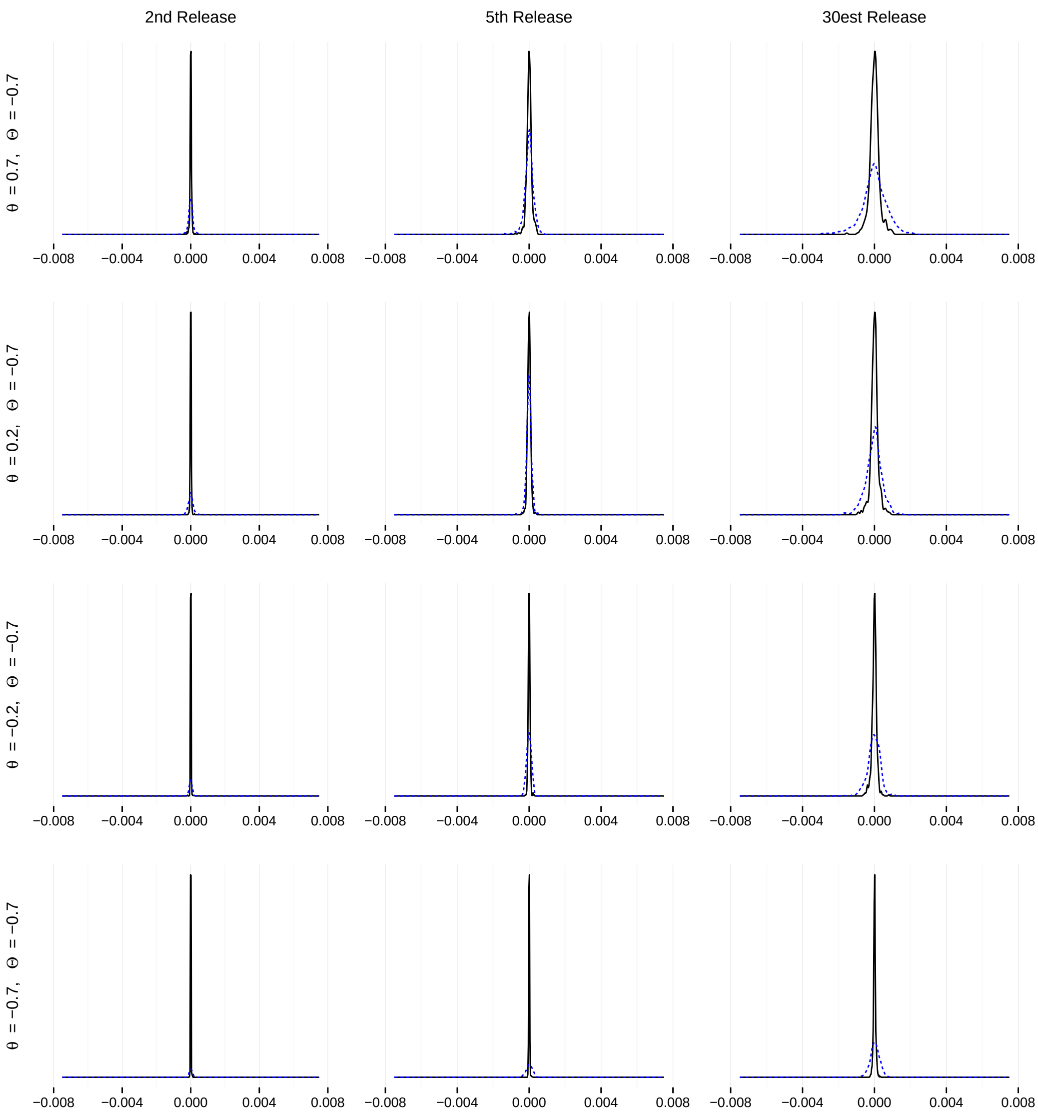
Also note that for larger negative values of Θ, the performance of X-11 is closer to SEATS (Figure 4.4). This may be due to the fact that these DGPs are closer to the Airline specification for which Depoutot and Planas (2002) found that the SEATS and X-11 filters are closest (see Section 4.32). To confirm this, Figure 4.5 shows the results of the same excercise but with the Depoutot-Planas model as a DGP. Accordingly, the estimated densities in this case are relatively close if the default specifications of both SEATS

Figure 4.1: Density Estimates for $\hat{\alpha}$. $\Theta = -0.2$



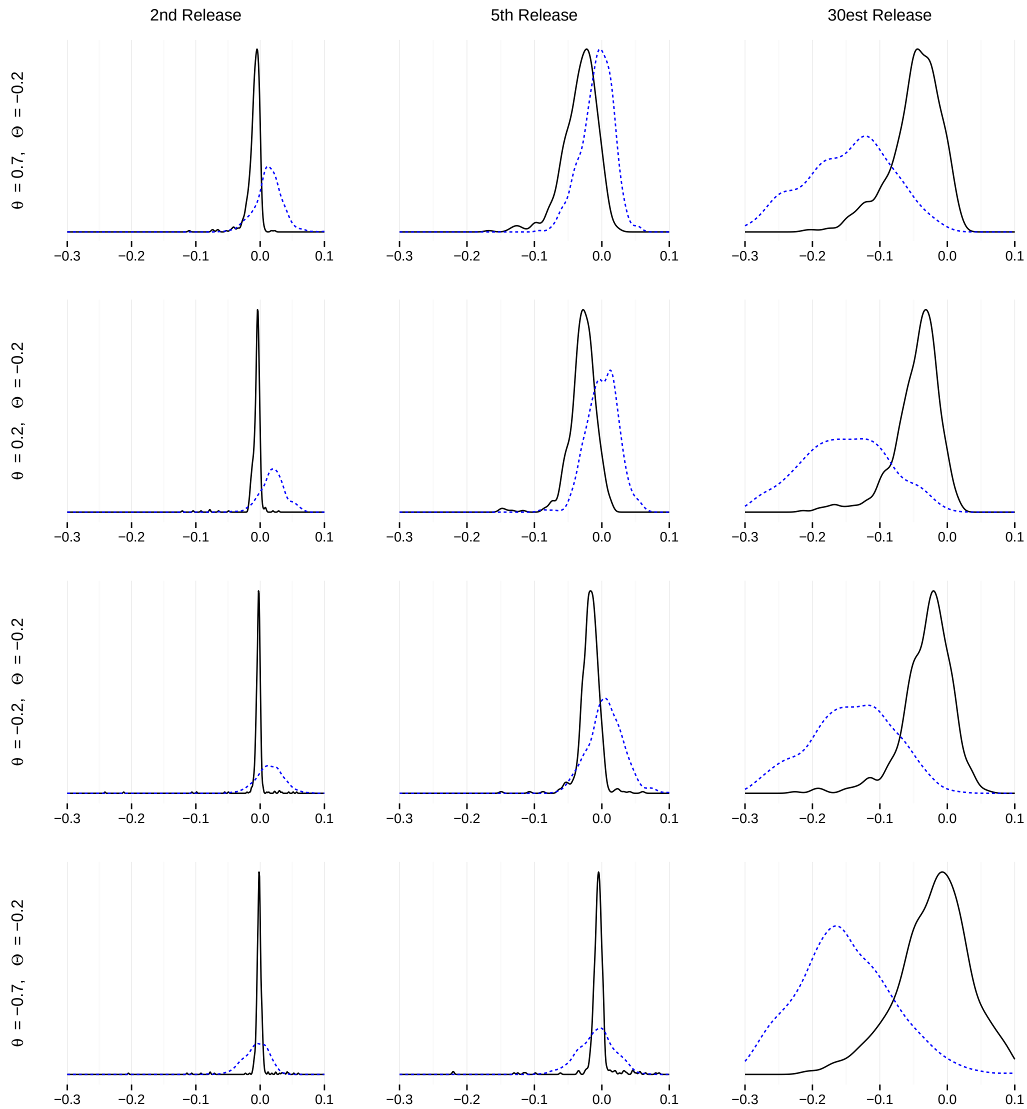
Note: Each graph shows density estimates of $\hat{\alpha}$ from the regression $rev_t^h = \alpha + \beta x_t^{initial} + \epsilon_t$. To calculate the densities, a Gaussian kernel estimator is used with the bandwidth choice following Silverman (1986)'s rule of thumb. Each density estimate is based on 500 $SARIMA(0,1,1)(0,1,1)$ series with $\{\theta, \Theta\} = \{0.7, -0.2\}$ for the first row, $\{\theta, \Theta\} = \{0.2, -0.2\}$ for the second row, $\{\theta, \Theta\} = \{-0.2, -0.2\}$ for the third row and $\{\theta, \Theta\} = \{-0.7, -0.2\}$ for the fourth row. The series were seasonally adjusted in real-time for $t = 140..500$, corresponding to about three decades of monthly data. Adjustment was conducted using both the X-11 (dotted lines) and X-13-SEATS (straight lines) algorithm available through the U.S. Census Bureau's X-13ARIMA-SEATS software. Revisions are calculated with respect to the second release (one revision), to the thirdiest release (29 revisions) and to the final release (number of revisions varies for each observation).

Figure 4.2: Density Estimates for $\hat{\alpha}$. $\Theta = -0.7$



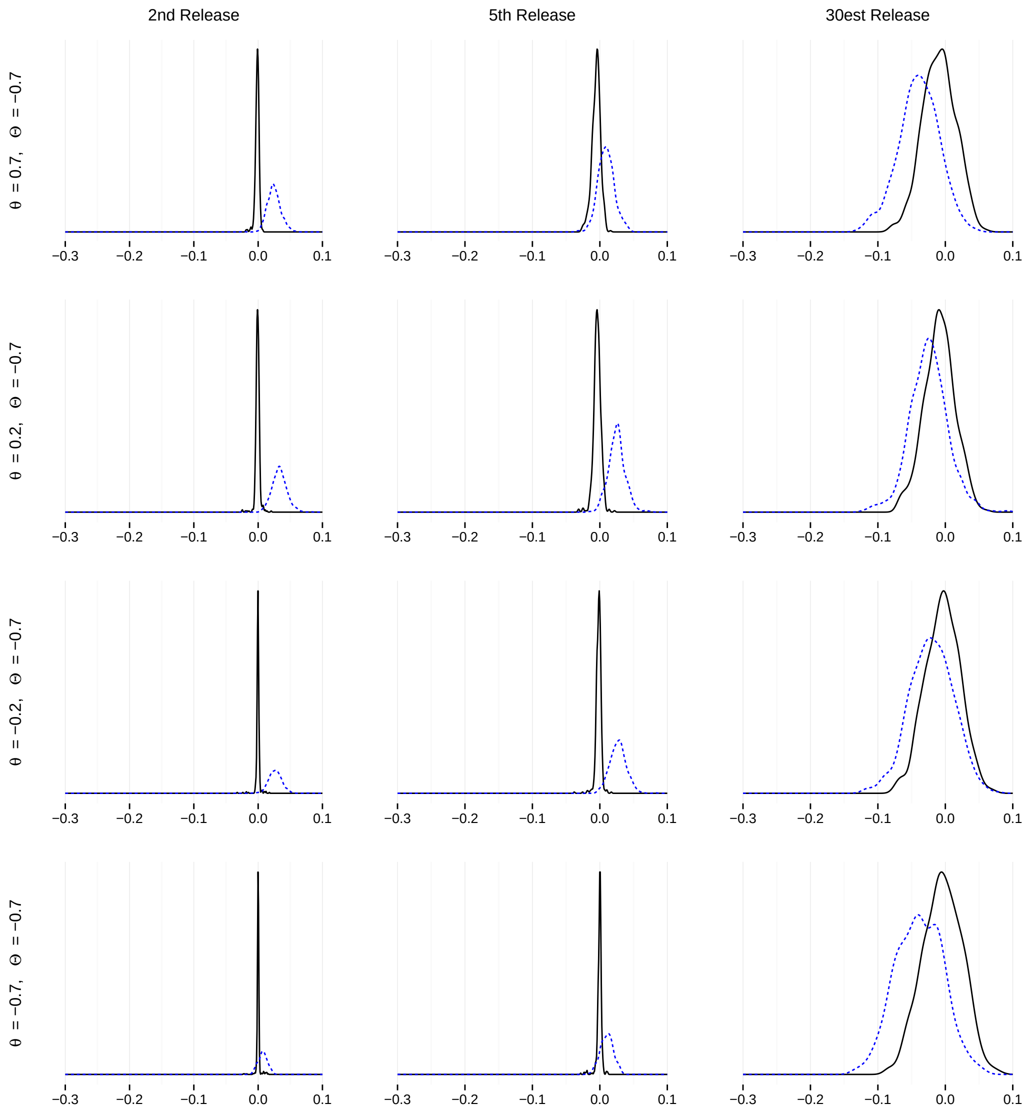
Note: Each graph shows density estimates of $\hat{\alpha}$ from the regression $rev_t^h = \alpha + \beta x_t^{initial} + \epsilon_t$. To calculate the densities, a Gaussian kernel estimator is used with the bandwidth choice following Silverman (1986)'s rule of thumb. Each density estimate is based on 500 $SARIMA(0,1,1)(0,1,1)$ series with $\{\theta, \Theta\} = \{0.7, -0.7\}$ for the first row, $\{\theta, \Theta\} = \{0.2, -0.7\}$ for the second row, $\{\theta, \Theta\} = \{-0.2, -0.7\}$ for the third row and $\{\theta, \Theta\} = \{-0.7, -0.7\}$ for the fourth row. The series were seasonally adjusted in real-time for $t = 140..500$, corresponding to about three decades of monthly data. Adjustment was conducted using both the X-11 (dotted lines) and X-13-SEATS (straight lines) algorithm available through the U.S. Census Bureau's X-13ARIMA-SEATS software. Revisions are calculated with respect to the second release (one revision), to the thirdiest release (29 revisions) and to the final release (number of revisions varies for each observation).

Figure 4.3: Density Estimates for $\hat{\beta}$. $\Theta = -0.2$



Note: Each graph shows density estimates of $\hat{\beta}$ from the regression $rev_t^h = \alpha + \beta x_t^{initial} + \epsilon_t$. To calculate the densities, a Gaussian kernel estimator is used with the bandwidth choice following Silverman (1986)'s rule of thumb. Each density estimate is based on 500 $SARIMA(0,1,1)(0,1,1)$ series with $\{\theta,\Theta\} = \{0.7,-0.2\}$ for the first row, $\{\theta,\Theta\} = \{0.2,-0.2\}$ for the second row, $\{\theta,\Theta\} = \{-0.2,-0.2\}$ for the third row and $\{\theta,\Theta\} = \{-0.7,-0.2\}$ for the fourth row. The series were seasonally adjusted in real-time for $t = 140..500$, corresponding to about three decades of monthly data. Adjustment was conducted using both the X-11 (dotted lines) and X-13-SEATS (straight lines) algorithm available through the U.S. Census Bureau's X-13ARIMA-SEATS software. Revisions are calculated with respect to the second release (one revision), to the thirdiest release (29 revisions) and to the final release (number of revisions varies for each observation).

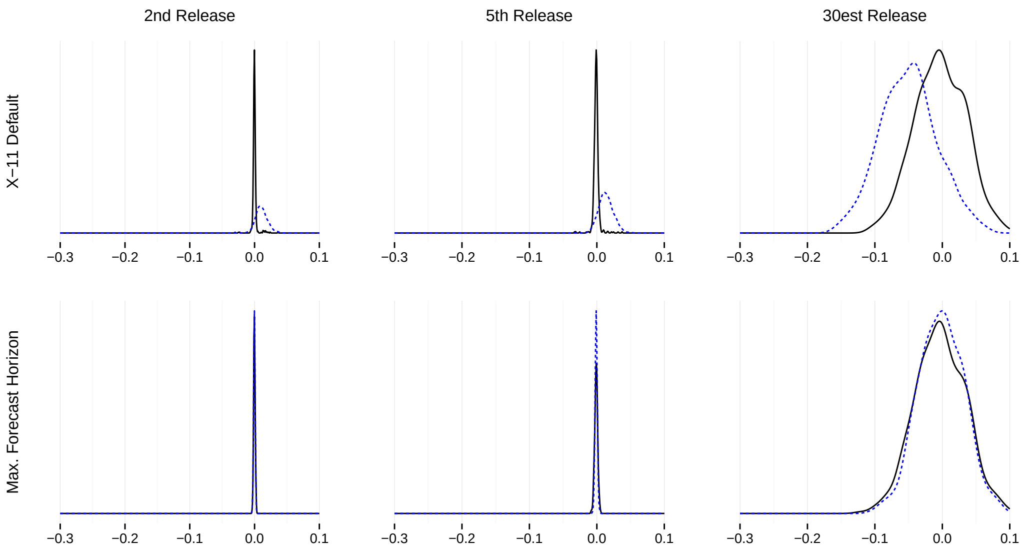
Figure 4.4: Density Estimates for $\hat{\beta}$. $\Theta = -0.7$



Note: Each graph shows density estimates of $\hat{\beta}$ from the regression $rev_t^h = \alpha + \beta x_t^{initial} + \epsilon_t$. To calculate the densities, a Gaussian kernel estimator is used with the bandwidth choice following Silverman (1986)'s rule of thumb. Each density estimate is based on 500 $SARIMA(0,1,1)(0,1,1)$ series with $\{\theta,\Theta\} = \{0.7,-0.7\}$ for the first row, $\{\theta,\Theta\} = \{0.2,-0.7\}$ for the second row, $\{\theta,\Theta\} = \{-0.2,-0.7\}$ for the third row and $\{\theta,\Theta\} = \{-0.7,-0.7\}$ for the fourth row. The series were seasonally adjusted in real-time for $t = 140..500$, corresponding to about three decades of monthly data. Adjustment was conducted using both the X-11 (dotted lines) and X-13-SEATS (straight lines) algorithm available through the U.S. Census Bureau's X-13ARIMA-SEATS software. Revisions are calculated with respect to the second release (one revision), to the thirdiest release (29 revisions) and to the final release (number of revisions varies for each observation).

and X-11 are used but model-based adjustment still outperforms in terms of forecast efficiency. However, for the second row of graphs in Figure 4.5 the forecast window in the preadjustement step (before the series are runned through the X-11 and the SEATS filters) was extended up to 5 years. In this case, both the densities of the estimates for β in Equation (4.2) as well as the size of the revisions (shown in the Appendix in Table 4.C.3) virtually coincide. This suggests that the asymmetric nature of X-11 at the end (and the beginning) of the time series is causing inefficiency in preliminary estimates of the seasonally adjusted data if the extension of the time series by forecasts is not long enough (at least for DGPs that are close to the Depoutot-Planas model).

Figure 4.5: Density Estimates for $\hat{\beta}$. $\theta = 0.583$, $\Theta = 0.551$ (Depoutot and Planas, 2002)



Note: See notes to Figure 4.4 and 4.3. Here, the DGP is an $SARIMA(0,1,1)(0,1,1)$ with $\{\theta, \Theta\} = \{-0.583, -0.551\}$. This corresponds to the Airline model which yields the SEATS (Wiener-Kolmogorov) filter that is closest to the default X-11 filter according to Depoutot and Planas (2002). The lower row of graphs uses a much longer forecast horizon in the preadjustment step for both SEATS and X-11 compared to the default specifications.

Note that the estimated distributions of $\hat{\beta}$ in the case of X-11 include the estimates found in Table 4.2 for Swiss unemployment and US employment and they are close to the range of estimates found by Faust, Rogers and Wright (2005). The latter use seasonally adjusted GDP data for G-7 countries to estimate Equation (4.2) and report β coefficients between -0.65 and -0.12 for long-term revisions. As they use quarterly data for about three decades, their setting corresponds most closely to the revisions with respect to the 30. or to the final release in the simulations above. Since the first versions of TRAMO-SEATS or comparable programs such as STAMP were developed no earlier than in the late eighties or early nineties, it is likely that one important reason for their findings is the use of moving-average type filters by the statistical agencies, most likely with a relatively short forecast window. In particular, prior to 1980 (and in many cases also later) the series were usually not extended by SARIMA-forecasts in the preadjustment step, before the actual adjustment was carried out.[18]

4.5 X-11 versus SEATS using Quarterly GDP Data

The results found in Section 4 confirm that seasonal adjustment is an important source for revisions to the adjusted data in general, depending on both the adjustment methods and the underlying DGP. In particular, these findings offer a potential explanation for the mean-reverting pattern of revisions for seasonally adjusted GDP found by Faust, Rogers and Wright (2005).

In this section, I roughly quantify to what extent mean-reverting GDP releases can be attributed to seasonal adjustment in the case of several countries. To do so I fit an $SARIMA(p,d,q)(P,D,Q)$ to real-world GDP data and then seasonally adjust the data in a pseudo real-time setting, just as in Section 4. For every adjustment, I use the default settings both for SEATS and X-11 but check for robustness using alternative choices. In particular, the results are not very different even if a much longer forecast horizon is used in the preadjustment step. To the best of my knowledge, no public real-time database for seasonally unadjusted GDP exists for OECD countries. This makes it impossible to isolate the effect of seasonal adjustment on revisions

[18] Note, however, that there may generally be a myriad of other reasons for revisions of GDP data. For example, in some cases statistical offices use growth rates of available data to forecast missing data. In this case, extreme growth rates of the available data will directly translate to extreme (high above/below their mean) growth rates of the missing data. As soon as the actual growth rates of the missing data becomes available, mean-reversion becomes very likely (I thank Jonathan Wright for this comment).

to GDP data since both the unadjusted data and the exact seasonal adjustment procedure used by the statistical agencies are unknown. However, in a pseudo real-time approach where all other sources of revisions are switched off and GDP series are directly seasonally adjusted, we can get a sense of how important this source of data uncertainty may be. Due to the fact that seasonal adjustment methods have improved with regard to both the size and the predictability of revisions (see below), these results may be interpreted as a lower bound for revisions caused by seasonal adjustment.

I focus on the countries from Faust, Rogers and Wright (2005) for which seasonally unadjusted data are available: France, Germany, Italy, Japan, and the UK. For Italy and Japan there are unfortunately no unadjusted data available for the earlier time period considered by Faust, Rogers and Wright (2005). This makes it more difficult to reconcile my results with theirs since seasonal factors may change over time and macroeconomic time series have generally become somewhat less volatile after 1990. However, I also provide additional evidence for Australia, Korea, Norway and Switzerland, where NSA data is available during a longer time period.

In a first step, I fit an SARIMA model with appropriate trading day and Easter effects to each GDP series using the latest available vintage. The model selection is based on the criterias used by statistical agencies for seasonal adjustment. Most notably, the resulting seasonally adjusted series should not exhibit any remaining seasonality and residuals should pass normality and autocorrelation tests.[19] For seasonal adjustment I assume that this model was known and used by the hypothetical statistical agency to adjust GDP for every vintage. I then calculate revisions to the first seasonally adjusted release with respect to the last vintage and re-run regression 4.2. The left panels of Table 4.1 show the results for X-11 adjustment (upper left panel) and the SEATS decomposition (lower left panel). The estimated coefficients for both α and β are depicted, where again a negative $\hat{\beta}$ provides evidence for mean-reversion. In addition, the R^2 and the Wald statistic, W, for joint significance of $\hat{\alpha}$ and $\hat{\beta}$ are provided. It appears indeed that X-

[19] X-13ARIMA-SEATS offers an automatic model selection procedure based on these and other criterias such as the AIC and BIC.

11 exhibits a somewhat stronger tendency to revert the preliminary release towards the mean: Almost all estimated coefficients for β are negative, in several cases significantly on the 10 % level. The results for SEATS on the other hand show parameter estimates that appear much more in-line with a distribution centered around zero, almost no significances and R^2 statistics that are slightly lower.

In a second step, I activate the automatic outlier detection of the X13-ARIMA-SEATS program to mimic the behavior of the agency to account for outliers, though in reality we may expect the agency to use more than only statistical information to detect and treat outliers. The results are displayed in the two middle panels of Table 4.1, again seperately for X-11 (top) and SEATS (bottom). Again, the findings point to the same direction, however, clearly with a more moderate outperformance by the model-based approach.

In a third step, I let X13-ARIMA-SEATS also fit the SARIMA model automatically for each vintage. This corresponds most closely to the situation of a statistical agency where the model choice is an important source of uncertainty.[20] Intuitively, model-based adjustment may be more sensitive to changes in the estimated model that is used for the decomposition, compared to the X-11 approach. However, the chosen model clearly also affects the latter since it matters a lot for the for- and backcasts at the edges of the time series, and hence for the estimated preliminary releases. The results are depicted in the right panels of Table 4.1. In this case, the outperformance by SEATS essentially disappears as both the estimated model and outliers change several times in case of most countries. Furthermore, it becomes also clear that seasonal adjustment in general, with standard settings and automatical outlier and model detection, can give rise to strikingly strong mean-reverting data revisions. The estimates are in the realm of the results found by Faust, Rogers and Wright (2005) for G7 GDP data during an earlier time period.

To reconcile this excercise with the simulation study above, I repeat step 1 and 2 using an Airline model. In most cases, this choice led to an accept-

[20] In reality, the models are often held fixed for several quarters before the statistical agency reevaluates its models and methods for seasonal adjustments.

Table 4.1: Mean Reversion in GDP Estimates Caused by Seasonal Adjustment

	X-11											
	Fix SARIMA				Fix SARIMA, auto outlier				Auto SARIMA, auto outlier			
	$\hat{\alpha}$	$\hat{\beta}$	R^2	W	$\hat{\alpha}$	$\hat{\beta}$	R^2	W	$\hat{\alpha}$	$\hat{\beta}$	R^2	W
UK	**0.08**	**-0.13**	0.07	**5.51**	**0.08**	**-0.11**	0.05	**6.22**	**0.09**	**-0.15**	0.09	**8.06**
JPN	0.04	-0.06	0.01	1.11	0.05	**-0.15**	0.15	**7.77**	0.06	-0.15	0.06	2.34
ITA	-0.01	**-0.18**	0.15	**12.6**	0.00	**-0.25**	0.26	**39.3**	-0.01	**-0.33**	0.31	**33.5**
GER	0.02	-0.03	0.00	0.66	0.02	-0.03	0.00	0.60	0.05	-0.09	0.02	2.60
FRA	0.02	**-0.04**	0.02	4.05	**0.02**	**-0.08**	0.06	**13.0**	**0.06**	**-0.14**	0.08	**10.2**
CHE	**0.12**	**-0.27**	0.19	3.72	**0.13**	**-0.30**	0.25	3.72	**0.17**	**-0.37**	0.30	**5.96**
KOR	0.12	-0.08	0.03	2.23	0.03	-0.03	0.00	0.41	-0.02	-0.00	0.00	1.01
AUS	-0.02	0.02	0.00	0.22	-0.03	0.06	0.00	1.69	**0.15**	**-0.18**	0.12	**9.86**
NOR	-0.02	0.02	0.00	0.89	0.00	0.01	0.00	0.34	-0.00	-0.00	0.00	0.03
	SEATS											
	Fix SARIMA				Fix SARIMA, auto outlier				Auto SARIMA, auto outlier			
	$\hat{\alpha}$	$\hat{\beta}$	R^2	W	$\hat{\alpha}$	$\hat{\beta}$	R^2	W	$\hat{\alpha}$	$\hat{\beta}$	R^2	W
UK	0.06	-0.11	0.05	2.49	0.07	-0.09	0.02	2.33	**0.07**	**-0.11**	0.05	**4.67**
JPN	0.02	0.04	0.00	1.70	**0.10**	**-0.16**	0.13	**6.34**	0.06	**-0.22**	0.09	3.83
ITA	-0.01	-0.17	0.10	**4.65**	0.00	**-0.27**	0.29	**19.8**	-0.01	**-0.44**	0.42	**34.1**
GER	-0.02	0.06	0.01	1.25	0.00	0.01	0.00	0.26	0.01	-0.00	0.00	0.12
FRA	-0.00	0.00	0.00	0.05	0.03	**-0.10**	0.08	**10.9**	**0.08**	**-0.16**	0.12	**10.4**
CHE	0.03	-0.08	0.02	1.07	0.09	-0.19	0.14	2.23	**0.17**	**-0.35**	0.29	**6.58**
KOR	0.03	-0.04	0.00	0.74	-0.08	0.02	0.00	1.46	-0.04	0.00	0.00	1.46
AUS	-0.03	0.04	0.00	0.93	-0.04	0.07	0.01	2.13	**0.24**	**-0.27**	0.24	**26.0**
NOR	-0.00	0.00	0.00	0.07	**0.03**	-0.00	0.00	3.75	-0.00	0.00	0.00	0.01

Note: The Table shows the results from regression 1. Bold values are statistically significant on the 10% level using Newey-West standard errors. For each regression, the R^2 and a Wald test for joint significance is also depicted. The two panels on the left show the results for X-11 (upper left) and SEATS (lower left) using an SARIMA fitted to the latest available vintage. This assumes that the latest vintage corresponds to the true data and that the statistical agency knew the model describing it from start. In the middle two panels, the automatical outlier detection of X-13ARIMA-SEATS was activated. In the right two panels, the automatical model detection was activated, i.e., the program was allowed to use another SARIMA model for every vintage.

able decomposition, that is, no seasonality was left in the adjusted series and standard criteria for the residuals were fullfilled (normality, no autocorrelation). Generally, the estimates for the parameters include $\hat{\theta} \in [-0.2, 0.1]$ and $\hat{\Theta} \in [-0.3, 0.7]$ which corresponds most closely to the results in the third row of Figure 4.C.5 and (to a lesser extent) to the third row of Figure 4.C.4. According to the findings in Figure 4.C.5, X-11 and SEATS should hence deliver relatively similar results with SEATS only mildly overperforming. As can be seen in Table 4.3, this is indeed what I find. if the automatic outlier detection is activated, the results are again comparable between the two approaches.

I reproduce all the above results but exclude large preliminary growth rates that may distort the results. To detect such outliers, I use a threshold

Table 4.2: Mean Reversion in GDP Estimates Caused by Seasonal Adjustment. Extreme Preliminary Growth Rates Deleted.

	X-11											
	Fix SARIMA				Fix SARIMA, auto outlier				Auto SARIMA, auto outlier			
	$\hat{\alpha}$	$\hat{\beta}$	R^2	W	$\hat{\alpha}$	$\hat{\beta}$	R^2	W	$\hat{\alpha}$	$\hat{\beta}$	R^2	W
UK	**0.11**	**-0.19**	0.08	**9.29**	**0.12**	**-0.19**	0.09	**9.78**	**0.11**	**-0.20**	0.08	**11.5**
JPN	**0.08**	-0.06	0.01	4.34	**0.07**	-0.08	0.02	3.02	**0.12**	**-0.18**	0.05	3.67
ITA	0.00	**-0.26**	0.18	**15.9**	0.00	**-0.26**	0.18	**15.9**	-0.05	**-0.22**	0.13	**25.0**
GER	0.04	-0.02	0.00	2.52	0.02	-0.02	0.00	0.98	0.04	-0.05	0.00	1.00
FRA	**0.03**	**-0.08**	0.03	**5.79**	0.01	-0.04	0.01	2.19	**0.07**	**-0.13**	0.05	**5.94**
CHE	0.05	-0.14	0.04	2.55	0.04	-0.13	0.04	2.82	**0.11**	**-0.26**	0.12	**7.85**
KOR	**0.17**	**-0.10**	0.02	3.97	0.09	-0.05	0.00	1.24	0.02	-0.01	0.00	0.10
AUS	0.03	0.00	0.00	2.84	0.03	0.02	0.00	3.35	**0.27**	**-0.30**	0.18	**28.4**
NOR	-0.01	0.00	0.00	0.57	0.01	-0.01	0.00	0.11	0.03	-0.03	0.00	0.34
	SEATS											
	Fix SARIMA				Fix SARIMA, auto outlier				Auto SARIMA, auto outlier			
	$\hat{\alpha}$	$\hat{\beta}$	R^2	W	$\hat{\alpha}$	$\hat{\beta}$	R^2	W	$\hat{\alpha}$	$\hat{\beta}$	R^2	W
UK	**0.07**	**-0.15**	0.05	4.30	0.03	-0.04	0.00	0.60	0.08	**-0.15**	0.04	3.53
JPN	0.06	0.02	0.00	**4.64**	**0.11**	-0.07	0.01	4.09	**0.15**	**-0.29**	0.07	**4.83**
ITA	-0.01	**-0.21**	0.12	**9.33**	0.00	**-0.28**	0.20	**7.08**	-0.04	**-0.29**	0.17	**11.2**
GER	-0.01	0.08	0.01	2.31	0.00	0.03	0.00	1.00	0.03	-0.02	0.00	0.84
FRA	0.00	-0.00	0.00	0.08	0.03	**-0.07**	0.02	2.88	**0.09**	**-0.15**	0.09	**10.6**
CHE	0.00	-0.01	0.00	0.14	-0.00	-0.00	0.00	0.36	**0.12**	**-0.23**	0.12	**6.42**
KOR	0.04	-0.02	0.00	0.18	0.05	-0.03	0.00	0.34	0.04	-0.03	0.00	0.28
AUS	0.01	0.00	0.00	1.18	0.00	0.02	0.00	1.93	**0.33**	**-0.33**	0.23	**30.7**
NOR	-0.00	-0.00	0.00	0.05	**0.03**	-0.01	0.00	3.67	0.01	-0.00	0.00	0.14

Note: See the note to Table 4.1.

of 2.5 times the standard deviation of preliminary GDP growth.[21] As can be seen in Tables 4.2 and 4.4, these results are entirely in-line with the previous findings.

Of course, all this has to be interpreted with caution since seasonal adjustment methods changed profoundly over time. First, many statistical agencies switched gradually from the original X-11 to the X-11-ARIMA and then the X-12-ARIMA program. Generally, these steps gave rise to better behaved revisions (Dagum and Laniel, 1987). Second, concurrent adjustment was introduced in some cases (e.g. by the Bureau of Labour Statistics in 2004), which is has been found to reduce the predictability of revisions (Kavajecz and Collins, 1995). In this study, I am using the latest implementation of both the X-11 and the SEATS algorithm and always concurrent adjustment. Therefore, my findings suggest that seasonal adjustment may have caused even stronger mean reversion in GDP data in the past when older methods

[21] Usually, between zero and five outliers were found using this definition.

Table 4.3: Mean Reversion in GDP Estimates Caused by Seasonal Adjsutment, Airline Model.

	X-11							
	Airline				Airline, auto outlier			
	$\hat{\alpha}$	$\hat{\beta}$	R^2	W	$\hat{\alpha}$	$\hat{\beta}$	R^2	W
UK	**0.08**	**-0.13**	0.07	**5.64**	**0.11**	**-0.17**	0.09	**14.4**
JPN	0.03	-0.08	0.04	0.82	0.04	-0.12	0.07	1.94
ITA	-0.01	**-0.16**	0.11	**8.58**	0.00	**-0.24**	0.23	**28.1**
GER	0.02	-0.03	0.00	0.53	0.01	-0.02	0.00	0.46
FRA	0.01	-0.02	0.00	0.88	0.02	**-0.05**	0.02	4.60
CHE	**0.12**	**-0.29**	0.21	4.14	**0.17**	**-0.39**	0.34	**4.70**
KOR	**0.22**	**-0.14**	0.09	4.02	0.24	-0.16	0.11	3.01
AUS	0.01	-0.01	0.00	0.12	0.00	0.01	0.00	0.67
NOR	0.00	0.00	0.00	0.08	0.03	-0.04	0.00	1.25
	SEATS							
	Airline				Airline, auto outlier			
	$\hat{\alpha}$	$\hat{\beta}$	R^2	W	$\hat{\alpha}$	$\hat{\beta}$	R^2	W
UK	0.07	-0.11	0.05	2.39	**0.10**	**-0.14**	0.05	**10.1**
JPN	0.02	-0.06	0.01	0.49	0.06	**-0.14**	0.07	3.26
ITA	-0.01	-0.16	0.09	3.90	0.01	**-0.27**	0.27	**17.4**
GER	-0.02	0.06	0.00	1.21	0.00	0.02	0.00	0.49
FRA	-0.00	0.02	0.00	0.94	0.01	**-0.03**	0.01	3.06
CHE	0.09	**-0.24**	0.14	3.27	**0.18**	**-0.42**	0.37	**5.35**
KOR	0.10	-0.08	0.01	1.58	0.14	-0.11	0.04	2.13
AUS	0.00	-0.00	0.00	0.18	-0.00	0.02	0.00	0.93
NOR	0.01	-0.02	0.00	1.03	0.02	-0.03	0.00	1.86

Note: See the note to Table 4.1. Here, the airline model was used, both without (left panels) and with automatical outlier detection (right panels).

were in use. Furthermore, GDP figures are usually calculated in a desaggregated fashion and therefore often indirectly seasonally adjusted. That is, each component of GDP is seasonally adjusted independently of all others and the adjusted components are then added up to produce the seasonally adjusted GDP figures. Finally, there are many different reasons why GDP data gets revised (see e.g. Indergand and Leist, 2014). It is hence very difficult to isolate what exactly causes predictability of revisions in real-world data.

Finally, the results in Section 4 suggest that different specifications than the default settings of X-13ARIMA-SEATS may improve the results. In particular, a much longer forecast horizon than the suggested default seems to improve the results, that is, revisions caused by the two seasonal adjustment algorithms may look much more similar. Nevertheless, the results in this section confirm that outlier and model uncertainty also play a very important

Table 4.4: Mean Reversion in GDP Estimates Caused by Seasonal Adjustment, Airline Model. Extreme Preliminary Growth Rates Deleted.

	X-11							
	Airline				Airline, auto outlier			
	$\hat{\alpha}$	$\hat{\beta}$	R^2	W	$\hat{\alpha}$	$\hat{\beta}$	R^2	W
UK	**0.11**	**-0.20**	0.08	**9.25**	**0.16**	**-0.27**	0.12	**12.3**
JPN	0.08	-0.15	0.07	2.77	0.11	**-0.22**	0.13	3.32
ITA	0.00	**-0.24**	0.16	**9.55**	0.00	**-0.24**	0.14	**7.97**
GER	0.04	-0.02	0.00	2.17	0.02	-0.00	0.00	0.83
FRA	0.03	-0.06	0.02	2.45	0.03	-0.07	0.02	2.09
CHE	0.06	**-0.17**	0.06	2.99	0.08	**-0.19**	0.08	3.75
KOR	**0.29**	**-0.17**	0.06	4.34	**0.34**	**-0.21**	0.09	2.98
AUS	0.06	-0.02	0.00	2.60	0.06	-0.02	0.00	2.84
NOR	-0.01	0.02	0.00	0.25	0.01	-0.03	0.00	0.57
	SEATS							
	Airline				Airline, auto outlier			
	$\hat{\alpha}$	$\hat{\beta}$	R^2	W	$\hat{\alpha}$	$\hat{\beta}$	R^2	W
UK	0.08	**-0.15**	0.05	3.77	0.04	-0.07	0.00	0.50
JPN	**0.09**	**-0.18**	0.07	**6.66**	**0.15**	**-0.35**	0.24	**12.4**
ITA	-0.01	**-0.17**	0.08	**5.84**	0.01	**-0.25**	0.16	**4.87**
GER	-0.01	0.07	0.01	2.33	0.00	0.04	0.00	1.27
FRA	0.01	-0.01	0.00	0.38	-0.00	-0.00	0.00	0.01
CHE	0.06	**-0.16**	0.06	4.20	0.09	**-0.18**	0.08	4.18
KOR	0.13	-0.09	0.01	1.61	**0.29**	**-0.18**	0.07	**5.11**
AUS	0.03	-0.02	0.00	0.79	0.06	-0.03	0.00	1.55
NOR	0.01	-0.02	0.00	0.67	0.02	-0.04	0.00	1.33

Note: See the note to Table 4.1. Here, the airline model was used, both without (left panels) and with automatical outlier detection (right panels).

role. While more research is warranted, this suggests that a careful modelling of the series in the preadjustement step is extremely important both for the model-based and moving-average based seasonal adjustment methods.

4.6 Conclusion

In this paper I show that preliminary releases of seasonally adjusted Swiss unemployment and US employment (household survey) figures are suboptimal in terms of forecast efficiency. The data exhibits similar, mean-reverting characteristics as the literature has found analyzing seasonally adjusted data for many variables in various countries. The problem originates to a large extent from the specific type of filter employed for seasonal adjustment. By simulation, and using the default settings of the new X-13ARIMA-SEATS software, I show that model-based filters (SEATS) deliver superior results compared to filters based on the X-11 algorithm.

First, the mean absolute revision to preliminary releases can be reduced by up to half if model-based seasonal adjustment methods are used. This is a substantial reduction, in particular in the light of the fact that policy makers generally rely on seasonally adjusted growth rates of economic variables - be it national accounts, the money stock or employment data - to form their forecasts and policy decisions.

Second, X-11 type filters tend to introduce a mean-reverting pattern into preliminary, seasonally adjusted releases. This means that large positive (negative) preliminary announcements tend to be revised downward (upward) in later releases. I find that this is a consistent feature over different data generating processes and becomes amplified during later releases. In case of the SEATS decomposition on the other hand, this pattern vanishes almost completely as long as the model is correctly specified and no large outliers distort the data. If model and outlier uncertainty are introduced, however, the performance of model-based adjustment also deteriorates.

These results help to explain mean-reversion of preliminary data found by recent studies, for example Faust, Rogers and Wright (2005) who analyze GDP of major economies, Garratt and Vahey (2006) who focus on many macroeconomic variables in the UK and Chapter 3 which analyzes Swiss national accounts variables. Using not seasonally adjusted GDP data for nine countries, I show in a pseudo real-time setting that seasonal adjustment accounts for this mean-reversion to a large extent.

For policy makers, this means that they should pay attention to the methods used by the statistical agencies to construct the data. Generally, preliminary releases generated from model-based seasonal filters seem to result in better behaved revisions compared to moving-average-based filters, assuming that both methods converge close to the true value. Hence, along the lines of Wright (2013), this paper suggests that statistical agencies should rely more on model-based seasonal adjustment. However, these filters are still comparatively new and developing further, and the X-11 family remains probably the most widely applied approach to date. This calls for further research.

Appendix

4.A A Brief Overview of the X-11 Algorithm

In the X-11 algorithm, the initial decomposition consists of the following steps. An initial trend, $y_t^{tr,init}$ is estimated using a centered 13-term moving average,

$$y_t^{tr,init} = \frac{1}{24}y_{t-6} + \frac{1}{12}y_{t-5} + .. + \frac{1}{12}y_t + .. + \frac{1}{12}y_{t+5} + \frac{1}{24}y_{t+6} \tag{4.14}$$

The resulting, initial detrended series (the so-called "SI-Ratio"),

$$SI_t^{init} = \frac{y_t}{y_t^{tr,init}}, \tag{4.15}$$

is then used to calculate a preliminary seasonal factor, $y_t^{s,pre}$, and an initial seasonal factor, $y_t^{s,init}$, by first applying a 3x3 seasonal moving average,

$$y_t^{s,pre} = \frac{1}{9}SI_{t-2s}^{init} + \frac{2}{9}SI_{t-s}^{init} + \frac{3}{9}SI_t^{init} + \frac{2}{9}SI_{t+s}^{init} + \frac{1}{9}SI_{t+2s}^{init} \tag{4.16}$$

and then a centered 13-term moving average:

$$y_t^{s,init} = \frac{y_t^{s,pre}}{\frac{1}{24}y_{t-6}^{s,pre} + \frac{1}{12}y_{t-5}^{s,pre} + .. + \frac{1}{12}y_t^{s,pre} + .. + \frac{1}{12}y_{t+5}^{s,pre} + \frac{1}{24}y_{t+6}^{s,pre}}. \tag{4.17}$$

The initial seasonally adjusted series then results as

$$y_t^{sa,init} = \frac{y_t}{y_t^{s,init}}. \tag{4.18}$$

Then the intermediate decomposition is carried out. This is done by applying a Henderson filter to $y_t^{sa,init}$ resulting in an intermediate trend estimate,

$$y_t^{tr,inter} = \sum_{j=-H}^{H} h_j^{(2H+1)} y_t^{sa,init}, h_j = h_{-j}, \tag{4.19}$$

where the Henderson weights are termined as outlined in the Appendices of Findley et al. (1998) and Wright (2013). Then a new detrended series, the intermediate SI-ratio, is calculated:

$$SI_t^{inter} = \frac{y_t}{y_t^{tr,inter}}, \tag{4.20}$$

The final seasonal factor, y_t^s, is calculated by applying a 3x5 seasonal moving average[22] and a centered 13-term moving average to SI_t^{inter} as follows:

$$\begin{aligned} y_t^{s,pre(2)} &= \frac{1}{15} SI_{t-3s}^{inter} + \frac{2}{15} SI_{t-2s}^{inter} + \frac{3}{15} SI_{t-s}^{inter} \\ &\quad + \frac{3}{15} SI_t^{inter} + \frac{3}{15} SI_{t+s}^{inter} + \frac{2}{15} SI_{t+s}^{inter} + \frac{1}{15} SI_{t+2s}^{inter} \end{aligned} \tag{4.21}$$

$$y_t^s = \frac{y_t^{s,pre(2)}}{\frac{1}{24} y_{t-6}^{s,pre(2)} + \frac{1}{12} y_{t-5}^{s,pre(2)} + .. + \frac{1}{12} y_t^{s,pre(2)} + .. + \frac{1}{12} y_{t+5}^{s,pre(2)} + \frac{1}{24} y_{t+6}^{s,pre(2)}}. \tag{4.22}$$

The "final" seasonally adjusted series, y_t^{sa}, is calculated as

$$y_t^{sa} = \frac{y_t}{y_t^s}. \tag{4.23}$$

The estimate for the "final" trend, y_t^t, is obtained by again applying a Henderson filter to y_t^{sa} and the "final" irregular component can then be

22 The X-11-ARIMA/88 and the X-12-ARIMA programs allow the user to choose among different seasonal moving average filters in this step.

calculated as

$$y_t^i = \frac{y_t^{sa}}{y_t^t}. \tag{4.24}$$

An outlier-adjustment algorithm is applied to the "final" irregular component where all extreme observations of a given month are replaced by the average value for this month over a 60-month window, yielding $y_t^{i,corr}$. See Wright (2013), p.105 for the definition of extreme observations in this context. Call the product of the "final" trend, the "final" seasonally adjusted series and the outlier-corrected "final" irregular series y_t^*,

$$y_t^* = y_t^t \times y_t^{sa} \times y_t^{i,corr}. \tag{4.25}$$

All the above steps are then repeated with y_t^* replacing the original data, y_t, where a potentially different (data-determined) Henderson-filter is used in Step 3.

Lastly, all steps are repeated a final time with a new extreme-value corrected y_t^{**}. Again, a potentially different Henderson-filter is used in Step 3 and, furthermore, a data-determined choice of a 3x3, 3x5 or 3x9 seasonal moving average in Equation (4.21). Again, see Wright (2013), pp.103-106 for details on the choice of the Henderson and the seasonal moving average filters in this context. Also note that I omit the trading day adjustment that may be applied in some circumstances to the irregular component prior to the extreme value correction.

4.B A Brief Overview of the SEATS Decomposition

In the model-based approach implemented in SEATS, the seasonal decomposition is viewed as a signal-extraction problem. Each of the unobservable components are assumed to follow a linear stochastic process. Hence, time series models for the components in Equation (4.9) have to be defined explicitly. These are derived from the linear model (4.11) used in the preadjustment steps. Of course, there exists no unique decomposition and, hence, several additional assumptions have to be made. The irregular component,

y_t^i, is assumed white noise whereas the trend component, y_t^t, and the seasonal component, y_t^s, are modelled as ARIMA processes,

$$\phi_t(L)(1-L)y_t^t = \theta_t(L)u_t^t, \tag{4.26}$$

and

$$\phi_s(L)(1+L+..+L^{s-1})y_t^s = \theta_s(L)u_t^s, \tag{4.27}$$

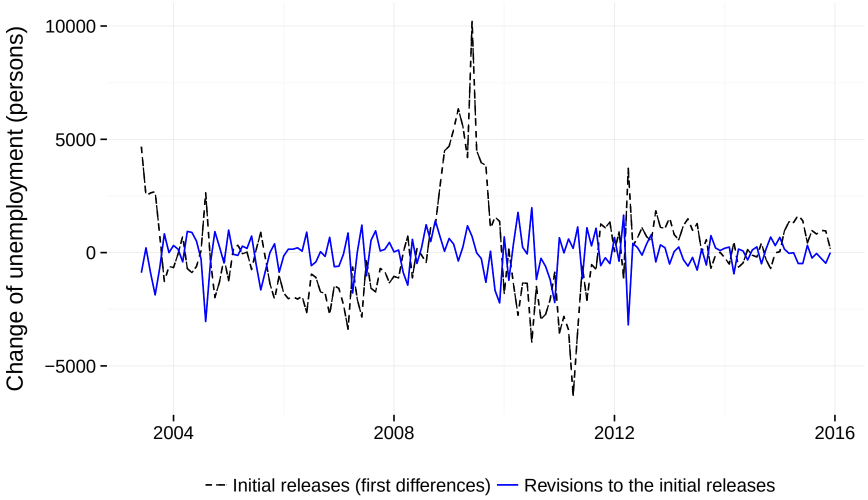
where both u_t^t and u_t^s are independently and identically distributed, $u_t^t \sim \mathcal{N}(0, \sigma_t^2)$ and $u_t^s \sim \mathcal{N}(0, \sigma_s^2)$. Several additional identifying assumtpions are necessary as outlined in Ghysels and Osborn (2001). Most importantly, y_t^t and y_t^s are assumed mutually independent. This is probably the strongest assumption as much has been written about common characteristics of the business and the seasonal cycle (see the overview in Miron, 1996). Furthermore, the decomposition with minimal variance of u_t^s is chosen among all admissible decompositions.

The estimation of the components can be done in two different ways which were shown to be equivalent Gómez (1999). The first is implemented in the STAMP program and involves a state space formulation of the component model in (4.26) and (4.27). The Kalman filter (for nonstationary series) can then be used to estimate the unobserved series. See Ghysels and Osborn (2001) and the references therein for more details.

The second strategy to estimate the components relies on the Wiener-Kolmogorov theory of signal extraction. The model is rewritten in terms of two components, where y_t^s is identified with the noise and the other components with the signal of interest. A Wiener-Kolmogorov-type filter can then be applied to the finite series extended with (infinite) forecasts and backcasts. In particular, this requires a partinioning of the autocovariance generating function (the spectral density in the frequency domain) of the observed series. For more details, see Gomez and Maravall (2001) and the references therein. For some more details about the implementation in X-13ARIMA-SEATS see Monsell (2007) and McElroy (2008).

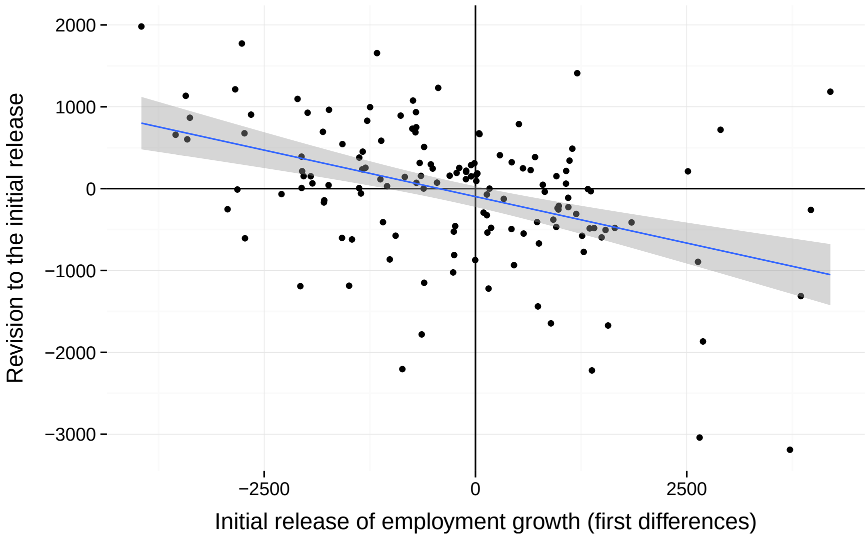
4.C Additional Figures and Tables

Figure 4.C.1: Swiss Unemployement Growth: Initial Releases and their Revisions



Note: Revisions are calculated based on the latest available vintage.

Figure 4.C.2: Swiss Registered Unemployment: Mean-Reversion of Initial Announcements



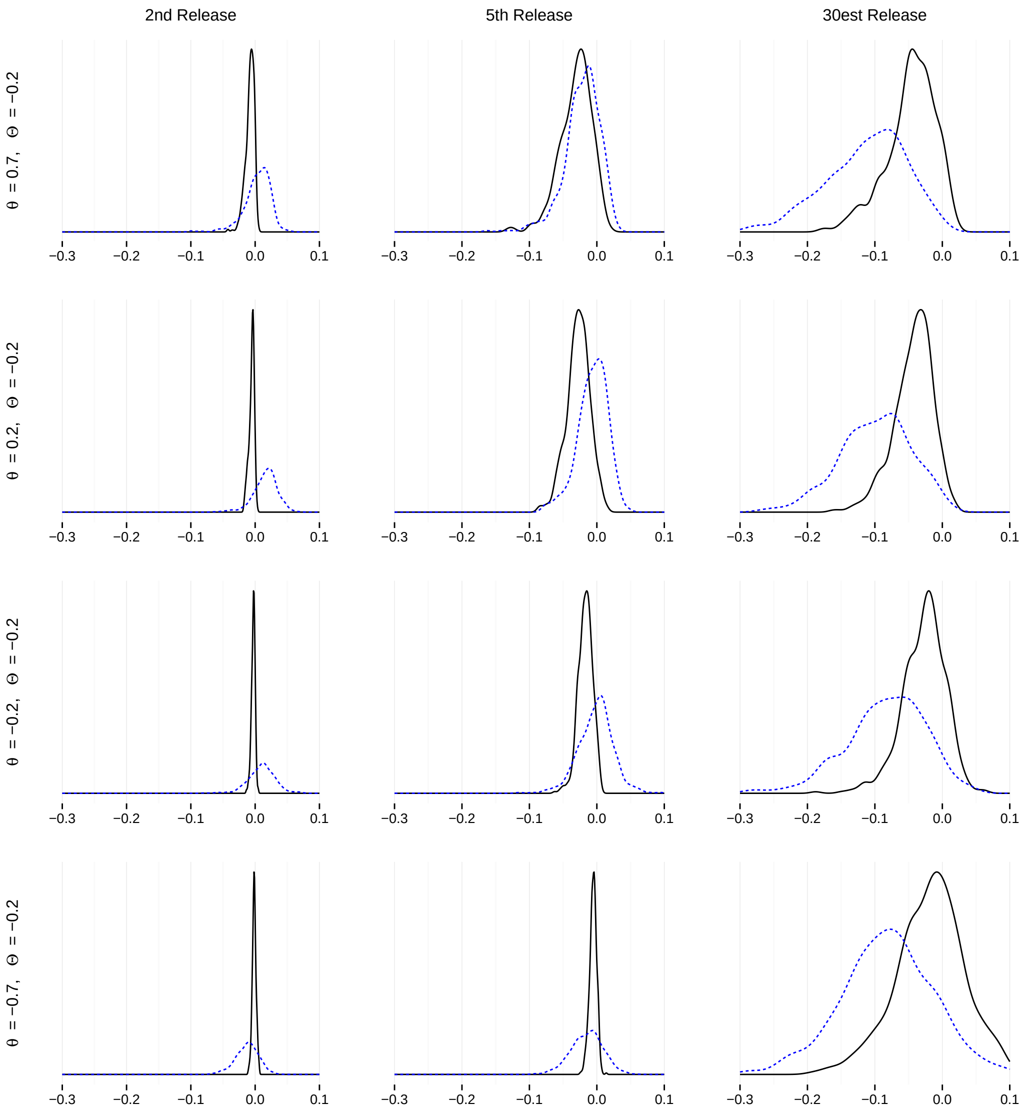
Note: Revisions are calculated based on the latest available vintage. The line depicts the beta coefficient in Equation (4.2) (see Table 4.2). The most extreme positive and negative first releases were omitted.

Table 4.C.1: Average Mean Revisions, X-11 with a 3x3 filter

	Average Mean Absolute Revision					
	2. Release		5. Release		30. Release	
	X-11	SEATS	X-11	SEATS	X-11	SEATS
$\{\theta, \Theta\} = \{0.7, -0.2\}$	0.150	0.076	0.188	0.148	0.435	0.263
$\{\theta, \Theta\} = \{0.2, -0.2\}$	0.118	0.039	0.142	0.098	0.348	0.222
$\{\theta, \Theta\} = \{-0.2, -0.2\}$	0.100	0.021	0.119	0.062	0.338	0.231
$\{\theta, \Theta\} = \{-0.7, -0.2\}$	0.075	0.015	0.096	0.029	0.394	0.244
$\{\theta, \Theta\} = \{0.7, -0.7\}$	0.119	0.032	0.146	0.064	0.327	0.246
$\{\theta, \Theta\} = \{0.2, -0.7\}$	0.100	0.021	0.120	0.045	0.275	0.209
$\{\theta, \Theta\} = \{-0.2, -0.7\}$	0.088	0.013	0.108	0.030	0.272	0.214
$\{\theta, \Theta\} = \{-0.7, -0.7\}$	0.067	0.007	0.088	0.016	0.313	0.257
	Average Mean Squared Revision					
	2. Release		5. Release		30. Release	
	X-11	SEATS	X-11	SEATS	X-11	SEATS
$\{\theta, \Theta\} = \{0.7, -0.2\}$	0.055	0.009	0.076	0.035	0.320	0.110
$\{\theta, \Theta\} = \{0.2, -0.2\}$	0.030	0.002	0.040	0.015	0.202	0.078
$\{\theta, \Theta\} = \{-0.2, -0.2\}$	0.023	0.001	0.028	0.006	0.194	0.084
$\{\theta, \Theta\} = \{-0.7, -0.2\}$	0.016	0.000	0.022	0.001	0.270	0.094
$\{\theta, \Theta\} = \{0.7, -0.7\}$	0.029	0.002	0.038	0.007	0.176	0.096
$\{\theta, \Theta\} = \{0.2, -0.7\}$	0.022	0.001	0.027	0.003	0.126	0.070
$\{\theta, \Theta\} = \{-0.2, -0.7\}$	0.020	0.000	0.025	0.002	0.127	0.073
$\{\theta, \Theta\} = \{-0.7, -0.7\}$	0.016	0.000	0.022	0.001	0.174	0.104

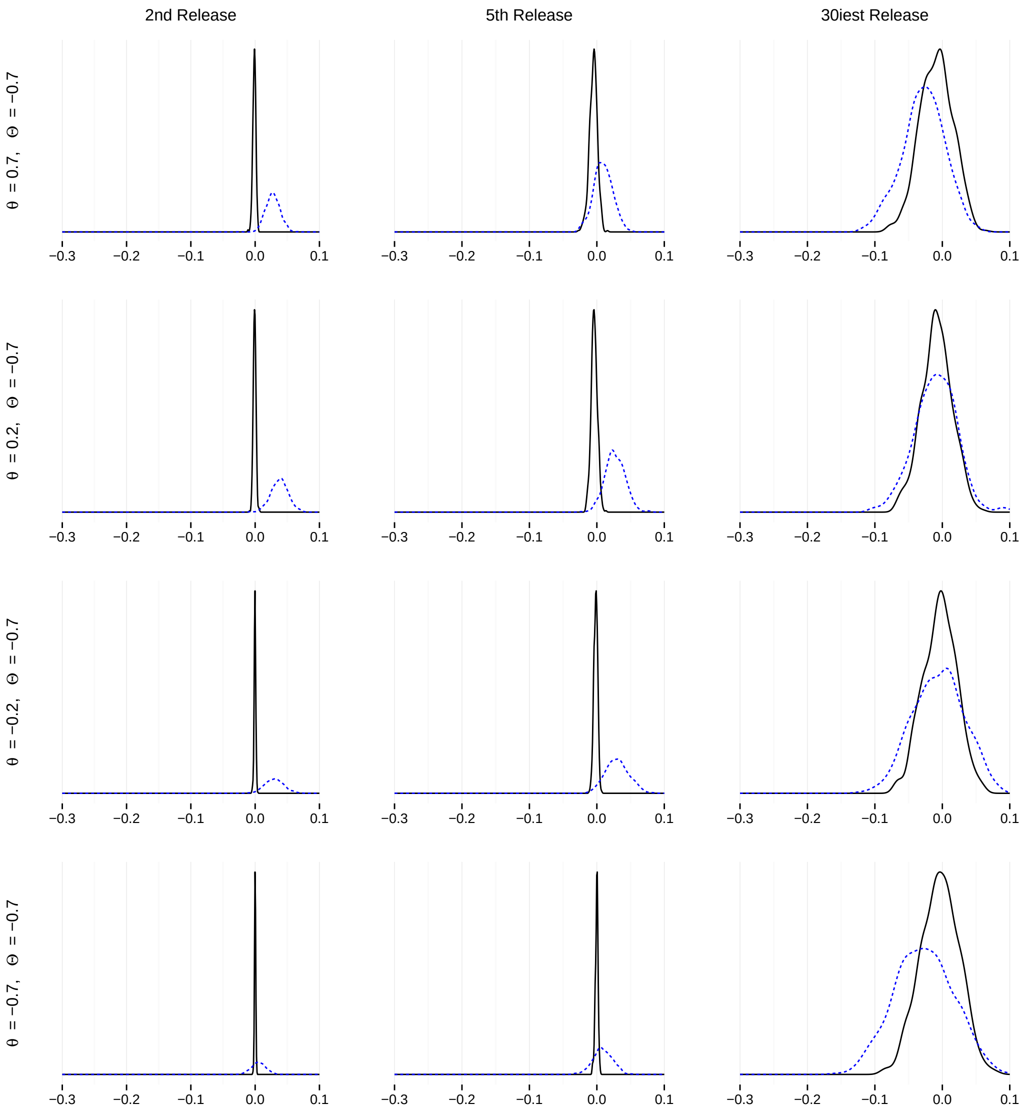
Note: The table shows mean absolute revisions and mean squared revisions in percentage points, averaged over 500 simulated series.

Figure 4.C.3: Density Estimates for $\hat{\beta}$. $\Theta = -0.2$, X-11 with a 3x3 filter



Note: Each graph shows density estimates of $\hat{\beta}$ from the regression $rev_t^h = \alpha + \beta x_t^{initial} + \epsilon_t$. To calculate the densities, a Gaussian kernel estimator is used with the bandwidth choice following Silverman (1986)'s rule of thumb. Each density estimate is based on 500 $SARIMA(0,1,1)(0,1,1)$ series with $\{\theta, \Theta\} = \{0.7, -0.2\}$ for the first row, $\{\theta, \Theta\} = \{0.2, -0.2\}$ for the second row, $\{\theta, \Theta\} = \{-0.2, -0.2\}$ for the third row and $\{\theta, \Theta\} = \{-0.7, -0.2\}$ for the fourth row. The series were seasonally adjusted in real-time for $t = 140..500$, corresponding to about three decades of monthly data. Adjustment was conducted using both the X-11 (dotted lines) and X-13-SEATS (straight lines) algorithm available through the U.S. Census Bureau's X-13ARIMA-SEATS software. Revisions are calculated with respect to the second release (one revision), to the thirdiest release (29 revisions) and to the final release (number of revisions varies for each observation).

Figure 4.C.4: Density Estimates for $\hat{\beta}$. $\Theta = -0.7$, X-11 with a 3x3 filter



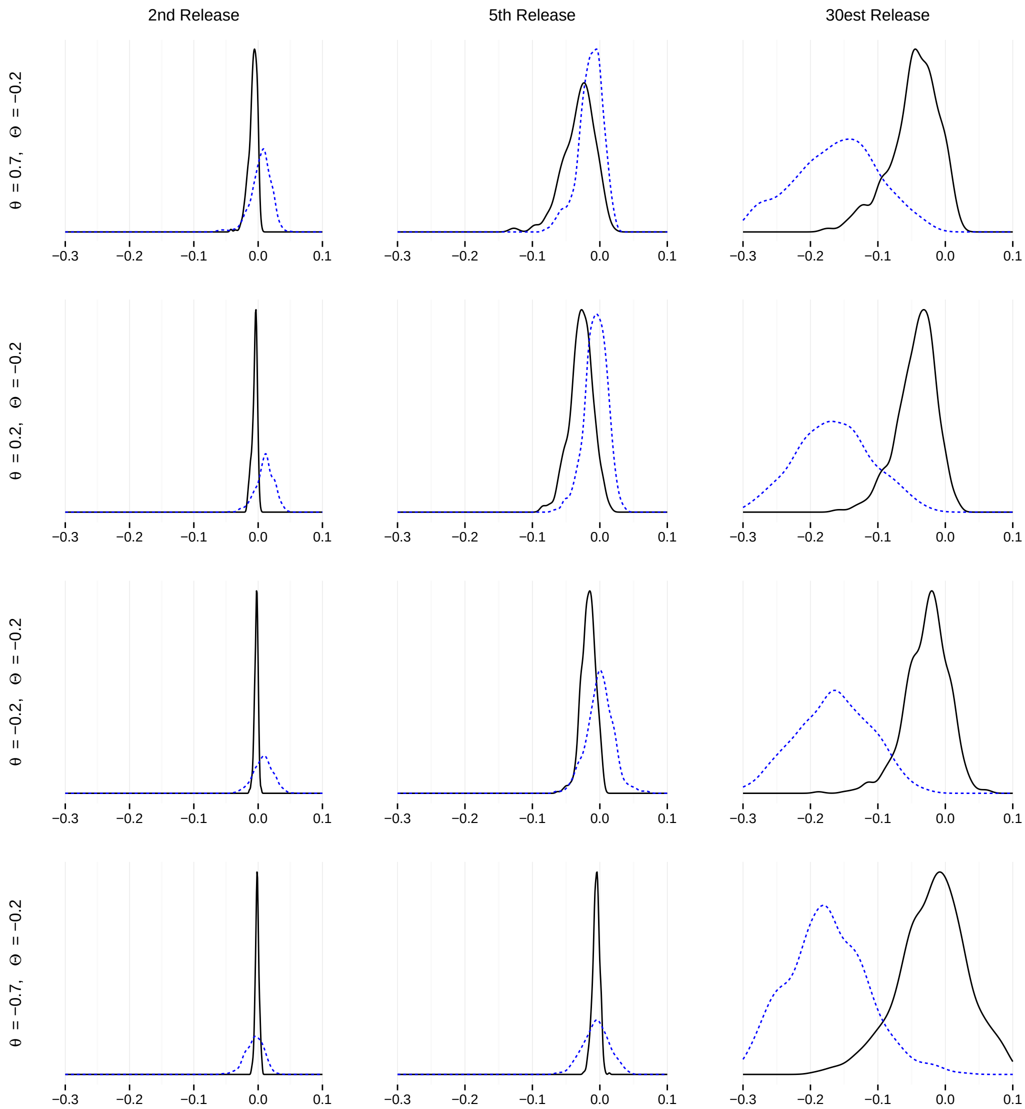
Note: Each graph shows density estimates of $\hat{\beta}$ from the regression $rev_t^h = \alpha + \beta x_t^{initial} + \epsilon_t$. To calculate the densities, a Gaussian kernel estimator is used with the bandwidth choice following Silverman (1986)'s rule of thumb. Each density estimate is based on 500 $SARIMA(0,1,1)(0,1,1)$ series with $\{\theta, \Theta\} = \{0.7, -0.7\}$ for the first row, $\{\theta, \Theta\} = \{0.2, -0.7\}$ for the second row, $\{\theta, \Theta\} = \{-0.2, -0.7\}$ for the third row and $\{\theta, \Theta\} = \{-0.7, -0.7\}$ for the fourth row. The series were seasonally adjusted in real-time for $t = 140..500$, corresponding to about three decades of monthly data. Adjustment was conducted using both the X-11 (dotted lines) and X-13-SEATS (straight lines) algorithm available through the U.S. Census Bureau's X-13ARIMA-SEATS software. Revisions are calculated with respect to the second release (one revision), to the thirdiest release (29 revisions) and to the final release (number of revisions varies for each observation).

Table 4.C.2: Average Mean Revisions, X-11 with a 3x5 filter

	Average Mean Absolute Revision					
	2. Release		5. Release		30. Release	
	X-11	SEATS	X-11	SEATS	X-11	SEATS
$\{\theta, \Theta\} = \{0.7, -0.2\}$	0.138	0.076	0.178	0.148	0.456	0.263
$\{\theta, \Theta\} = \{0.2, -0.2\}$	0.107	0.039	0.134	0.098	0.370	0.222
$\{\theta, \Theta\} = \{-0.2, -0.2\}$	0.092	0.021	0.113	0.062	0.368	0.231
$\{\theta, \Theta\} = \{-0.7, -0.2\}$	0.073	0.015	0.094	0.029	0.442	0.244
$\{\theta, \Theta\} = \{0.7, -0.7\}$	0.087	0.032	0.108	0.064	0.294	0.246
$\{\theta, \Theta\} = \{0.2, -0.7\}$	0.075	0.021	0.091	0.045	0.250	0.209
$\{\theta, \Theta\} = \{-0.2, -0.7\}$	0.068	0.013	0.083	0.030	0.254	0.214
$\{\theta, \Theta\} = \{-0.7, -0.7\}$	0.051	0.007	0.069	0.016	0.301	0.257
	Average Mean Squared Revision					
	2. Release		5. Release		30. Release	
	X-11	SEATS	X-11	SEATS	X-11	SEATS
$\{\theta, \Theta\} = \{0.7, -0.2\}$	0.055	0.009	0.075	0.035	0.350	0.110
$\{\theta, \Theta\} = \{0.2, -0.2\}$	0.028	0.002	0.039	0.015	0.226	0.078
$\{\theta, \Theta\} = \{-0.2, -0.2\}$	0.021	0.001	0.028	0.006	0.224	0.084
$\{\theta, \Theta\} = \{-0.7, -0.2\}$	0.016	0.000	0.022	0.001	0.324	0.094
$\{\theta, \Theta\} = \{0.7, -0.7\}$	0.015	0.002	0.021	0.007	0.140	0.096
$\{\theta, \Theta\} = \{0.2, -0.7\}$	0.012	0.001	0.016	0.003	0.102	0.070
$\{\theta, \Theta\} = \{-0.2, -0.7\}$	0.011	0.000	0.014	0.002	0.107	0.073
$\{\theta, \Theta\} = \{-0.7, -0.7\}$	0.008	0.000	0.012	0.001	0.153	0.104

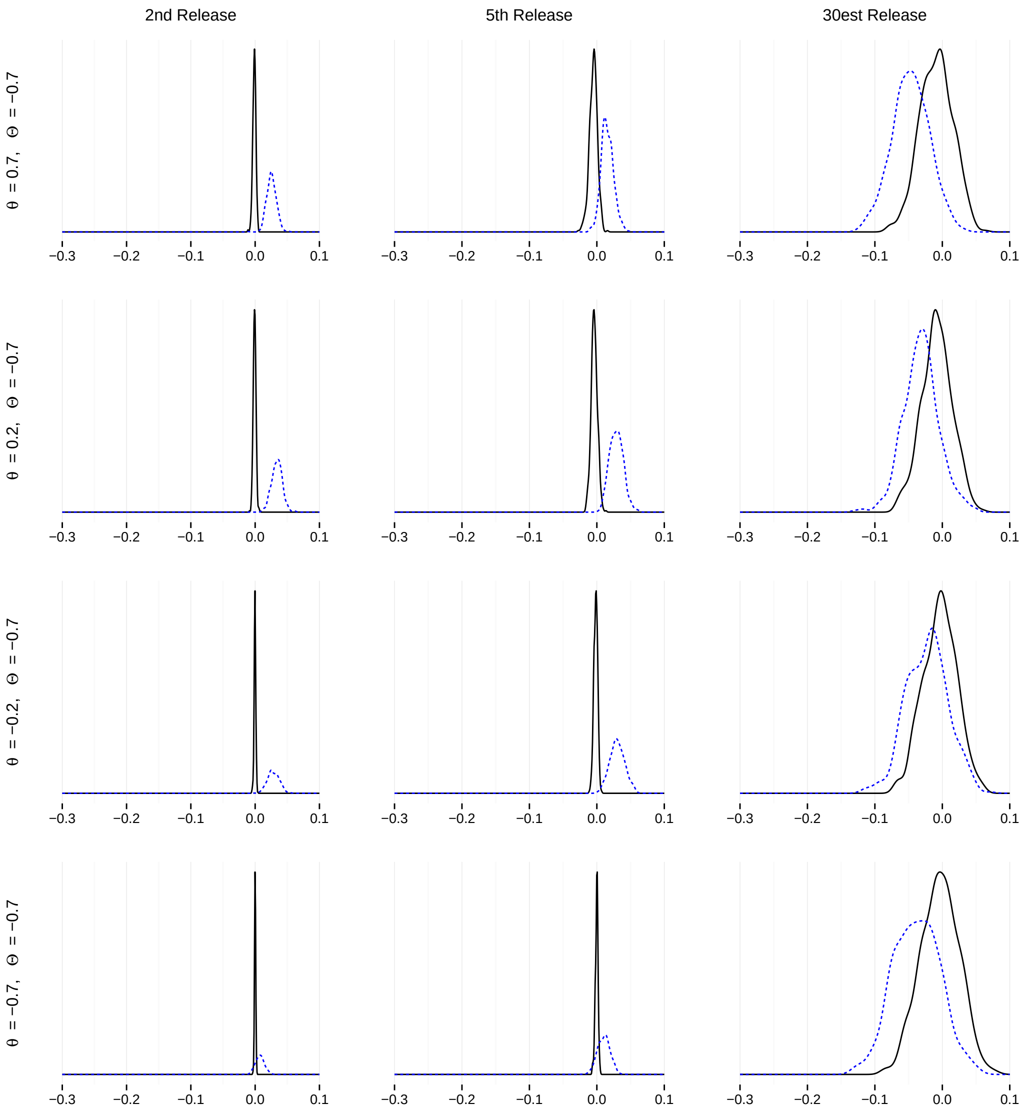
Note: The table shows mean absolute revisions and mean squared revisions in percentage points, averaged over 500 simulated series.

Figure 4.C.5: Density Estimates for $\hat{\beta}$. $\Theta = -0.2$, X-11 with a 3x5 filter



Note: Each graph shows density estimates of $\hat{\beta}$ from the regression $rev_t^h = \alpha + \beta x_t^{initial} + \epsilon_t$. To calculate the densities, a Gaussian kernel estimator is used with the bandwidth choice following Silverman (1986)'s rule of thumb. Each density estimate is based on 500 $SARIMA(0,1,1)(0,1,1)$ series with $\{\theta, \Theta\} = \{0.7, -0.2\}$ for the first row, $\{\theta, \Theta\} = \{0.2, -0.2\}$ for the second row, $\{\theta, \Theta\} = \{-0.2, -0.2\}$ for the third row and $\{\theta, \Theta\} = \{-0.7, -0.2\}$ for the fourth row. The series were seasonally adjusted in real-time for $t = 140..500$, corresponding to about three decades of monthly data. Adjustment was conducted using both the X-11 (dotted lines) and X-13-SEATS (straight lines) algorithm available through the U.S. Census Bureau's X-13ARIMA-SEATS software. Revisions are calculated with respect to the second release (one revision), to the thirdiest release (29 revisions) and to the final release (number of revisions varies for each observation).

Figure 4.C.6: Density Estimates for $\hat{\beta}$. $\Theta = -0.7$, X-11 with a 3x5 filter



Note: Each graph shows density estimates of $\hat{\beta}$ from the regression $rev_t^h = \alpha + \beta x_t^{initial} + \epsilon_t$. To calculate the densities, a Gaussian kernel estimator is used with the bandwidth choice following Silverman (1986)'s rule of thumb. Each density estimate is based on 500 $SARIMA(0,1,1)(0,1,1)$ series with $\{\theta, \Theta\} = \{0.7, -0.7\}$ for the first row, $\{\theta, \Theta\} = \{0.2, -0.7\}$ for the second row, $\{\theta, \Theta\} = \{-0.2, -0.7\}$ for the third row and $\{\theta, \Theta\} = \{-0.7, -0.7\}$ for the fourth row. The series were seasonally adjusted in real-time for $t = 140..500$, corresponding to about three decades of monthly data. Adjustment was conducted using both the X-11 (dotted lines) and X-13-SEATS (straight lines) algorithm available through the U.S. Census Bureau's X-13ARIMA-SEATS software. Revisions are calculated with respect to the second release (one revision), to the thirdiest release (29 revisions) and to the final release (number of revisions varies for each observation).

Table 4.C.3: Average Mean Revisions for $\theta = 0.583$, $\Theta = 0.551$ (Depoutot and Planas, 2002)

	Average Mean Absolute Revision					
	2. Release		5. Release		30. Release	
	X-11	SEATS	X-11	SEATS	X-11	SEATS
Default Specifications	0.058	0.012	0.075	0.025	0.322	0.309
Max. Forecast Window	0.011	0.011	0.023	0.023	0.285	0.285
	Average Mean Squared Revision					
	2. Release		5. Release		30. Release	
	X-11	SEATS	X-11	SEATS	X-11	SEATS
Default Specifications	0.010	0.000	0.013	0.001	0.173	0.150
Max. Forecast Window	0.001	0.000	0.001	0.001	0.128	0.128

Note: The table shows mean absolute revisions and mean squared revisions in percentage points, averaged over 500 simulated series.

Chapter 5

Which Factors Drive the Skill-Mix of Migrants in the Long-Run?

Andreas Beerli, Ronald Indergand

Summary: Immigrants in OECD countries have become increasingly highly skilled and more positively selected between 1980 and 2010. This paper analyses the factors driving these trends focusing on Switzerland, which experienced large immigration inflows and similar trends. Our findings suggest that skill-biased demand shifts in destinations and education upgrading in the origin countries of immigrants are the most important determinants. Yet, while education upgrading predicts only a modest increase in tertiary and a larger increase in secondary education, we show that skill-biased demand shifts are crucial to understand why tertiary education increased sharply among immigrants while the share of secondary educated merely stabilized. Additionally, we show that our results are not driven by different migration policy regimes.

We thank David Autor, David Dorn, Georg Graetz, Jeffrey Grogger, Alan Manning, Klaus Neusser, Gianmarco Ottaviano, Giovanni Peri, Kevin Shih, Michael Siegenthaler, Rainer Winkelmann, Ulrich Woitek, Fabrizio Zilibotti and Joseph Zweimüller and seminar participants at the University of Zurich, the University of Bern, the London School of Economics and Political Science, the University of California in Davis, the European Economic Association Meeting and the University of Pennsylvania for valuable comments and discussions.

5.1 Introduction

International migrants in developed countries are increasingly highly skilled. Between 1980 and 2010, the share of immigrants with a tertiary education increased by 19 percentage points on average across 19 OECD destination countries.[1] Meanwhile, the share of immigrants with secondary schooling increased by 8 percentage points and the share with no or primary schooling fell by 27 percentage points.

An interesting exercise is to contrast these trends in the skill-mix of immigrants with the trends in their home countries. Each panel in Figure 5.1 plots the share of an education group among immigrants in OECD countries against the corresponding share in their home country populations, both for 1980 (gray lower case letters) and 2010 (red upper case letters). Panel A shows the result for tertiary (highly) educated, Panel B for secondary (middle) educated and Panel C for primary or less educated (low) immigrants.[2] Each panel also shows the 45-degree line and the cross-destination mean. Figure 5.1 shows that immigrants were positively selected in most OECD countries already in 1980: their share of tertiary educated was larger compared to their origin country population (17% vs. 8% on average). In contrast, middle educated individuals were underrepresented among immigrants in 1980 (24% vs. 34% on average) whereas the share of low educated was almost equal (59% vs. 58%). While this pattern of *positive selection* has been widely acknowledged in the literature (e.g. see discussion in Bertoli et al., 2012), its *change between 1980 and 2010* has received, to our knowl-

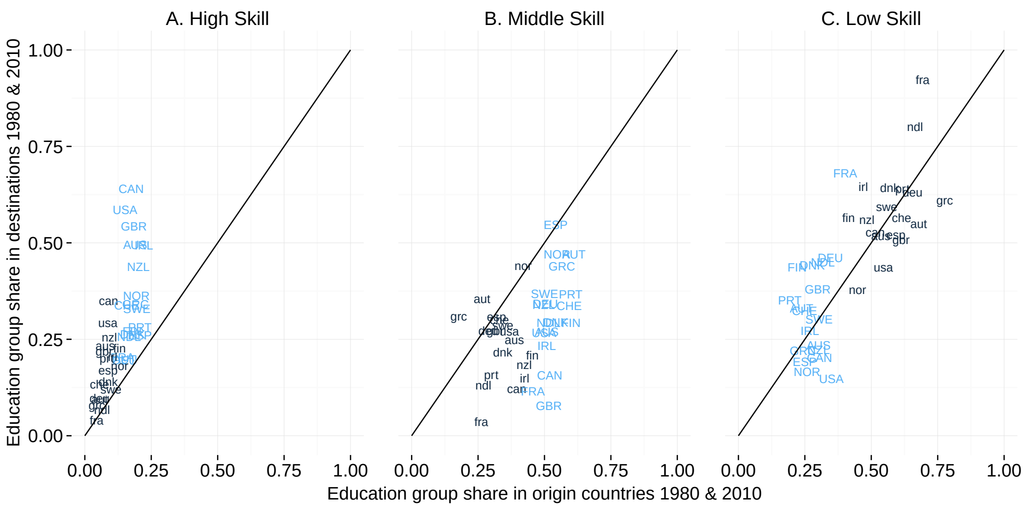
[1] Data from Brücker, Capuano and Marfouk (2013). The countries are Australia, Austria, Canada, Chile, Denmark, Finland, France, Germany, Greece, Ireland, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and United States. We drop Luxembourg and Chile and replace the values for Switzerland with values from the Swiss Census 1980 and 2010.

[2] More specifically, we use the share of immigrants from an origin country o on the total immigrant population in a destination j, $m_{j,o,t}$ from the Brücker, Capuano and Marfouk (2013) data and the education shares in their home countries from Barro and Lee (2013), $EDUSH^{e}_{o,t}$, to construct a weighted average education share of the home country populations of immigrants in a destination, $\widetilde{EDUSH}^{e}_{j,t} = \sum_o m_{j,o,t} \cdot EDUSH^{e}_{o,t}$ where $e \in \{\text{high, middle, low}\}$. These data sets use the same education definitions. Cf. Section 3 for more information about the data sets used.

edge, almost no attention. A first contribution of this paper is to document this change in the selection of migrants.

Between 1980 and 2010, the share of highly educated and middle educated increased both among immigrants and in the home countries, but highly educated individuals became even more over-represented among immigrants (36% vs. 18% on average) while middle educated individuals became more under-represented (32% vs. 58% on average). Thus, the positive selection of highly educated relative to middle educated *accentuated* considerably. In contrast, the share of low educated among immigrants decreased almost in lock-step with the corresponding group share in the home countries (to 32% vs 29% on average). The figure also shows that there is substantial variation across destination countries with large gains of highly educated immigrants in some English-speaking countries in contrast to more modest increases in many continental European countries.

Figure 5.1: Skill Group Shares of Immigrants in 18 OECD Destination Countries and Average Skill Group Shares in the Origin Country Populations, 1980 and 2010



Note: The vertical axis shows the share of the particular education group among foreign-born in each destination country in 1980 (gray lower case letters) and 2010 (red upper case letters) based on data from Brücker, Capuano and Marfouk (2013). The horizontal axis shows for each destination country the average share of that education group in the origin countries based on data from Barro and Lee (2013) taking population size of an origin country in a destination as weight. The solid bar represents a 45-degree line. ‘‘mean1980’’ (‘‘Mean2010’’) represent the average education group share across destinations in 1980 and 2010, respectively. Highly educated workers have a some tertiary education, middle educated workers have some secondary education and low educated workers have compulsory schooling or less.

These trends have gained more saliency as highly educated immigrants are often seen as net-contributors to destination countries' economies (Dustmann and Frattini, 2014), potentially fostering innovation and productivity (Kerr and Lincoln, 2010; Shih, Peri and Sparber, 2015). The public discussion revolves around the question of how governments may ‘‘manage’’ immigration to alleviate shortages of skills (Chaloff and Lemaitre, 2009; Stevens, di Mattia and Schieb, 2009) with some heralding that the ‘‘battle for the brains‘‘ is about to unfold (Bertoli et al., 2012). Yet, there is little agreement on the actual drivers of skill shortage and whether and how policy makers ought to respond.

In this paper, we analyze the determinants of these trends by focusing on newly arriving immigrants in Switzerland from 1980 to 2010. We use variation in the skill-mix of new immigrants from 30 different origin countries between 1980 and 2010 across Swiss local labor markets. Switzerland is an interesting laboratory to study immigration for a number of reasons. First, together with a group of other countries such as Australia, Canada and the U.S., Switzerland has exhibited very high immigration rates. In 2010, the country's population share of foreign born was at 29% (up from 17% in 1980), only surpassed by Luxembourg (OECD, 2014). Second, Switzerland has witnessed a strong change in the skill composition of immigrants paralleling the international experience depicted in Figure 5.1: The positive selection of highly over middle educated immigrants accentuated considerably between 1980 and 2010. Third, exploiting variation across many local labor markets in Switzerland, where new immigrants arrive and settle for the first time, and across a large set of origin countries allows disentangling different forces as we explain next.

To study the determinants of the trends, we build on a Roy model of immigrant selection adapted to education groups by Grogger and Hanson (2011), henceforth GH. According to this framework four factors determine the skill-composition of immigrants in a destination: (i) the earnings difference between workers with different education in the destination, (ii) the earnings differences in the origin country, (iii) the population shares of the education groups in the origin country and (iv) education specific bilateral migration costs. Figure 5.1 shows intriguingly that the change in the skill structure in the origin countries of immigrants, i.e. factor (iii), may explain the strong decrease in the share of low educated immigrants but falls short of explaining the trends favoring highly over middle educated immigrants. Hence, which other forces could have contributed to this trend?

The existing literature offers no conclusive explanation. An emerging result from the literature on immigrant selection is that skill-premiums in destination countries, i.e. factor (i), conditional the skill-premiums in origin countries, i.e. factor (ii), play an important role for immigrant selection and sorting.[3] Based on a cross-sectional analysis, GH, for instance, find that the skill-premium that earnings differences of highly relative to low educated in destinations predicts well the relative population share of these two skill groups among migrants. Bertoli et al. (2012) analyze the selection of highly relative to low skilled in a cross-country panel and find that also the adoption of selective migration policies favoring highly skilled, i.e. factor (iv), plays a role. Both studies, however, are essentially silent on the underlying cause of the trends favoring highly over middle educated migrants between 1980 and 2010. Additionally, both studies treat skill-premiums as exogenous despite a large literature discussing the impact of immigrants on the wage distribution.[4] Another strand of the literature discusses how the skill-premiums of highly relative to middle skilled workers increased during the last 30 years in most OECD countries due to the forces of technology and globalization (Acemoglu and Autor, 2011; Goos, Manning and Salomons, 2014). In particular, highly educated workers have experienced an increase in the demand for their labor, raising both their employment and wages while the employ-

[3] Some authors highlight that additional factors might modulate the effect of differences in the returns to skills between destination and home country. E.g. Chiquiar and Hanson (2005) point out that lower migration costs for higher educated could explain why Mexican immigrants in the US are selected from the middle rather than low levels of education.McKenzie and Rapoport (2010) show that immigrants are more negatively selected if they had access to close migration networks. Additionally, Belot and Hatton (2012) point out that return to education becomes only important after conditioning on poverty constraints.

[4] Dustmann, Frattini and Preston (2013), for instance, find heterogenous effects along the distribution of native wages. Bertoli et al. (2012) acknowledge the potential endogeneity of the adoption of selective immigration policies and provide a discussion.

ment and the wages of middle skilled workers in medium paying jobs declined relatively. This literature, however, is silent on whether and how these forces affected international migrants.[5]

This paper makes a contribution at the intersection of these two strands of the literature by analyzing the factors driving the trends in the skill-mix of migrants with specific emphasis on the role of skill- or routine-biased technical change. Using the variation in the skill-mix of new immigrants from different origin countries across different regions in Switzerland allows exploiting the fact that these regions were exposed to different degrees of routine-biased demand shifts due to their pre-existing industrial structure. In particular, building on Autor and Dorn (2013), we use the fact that regions with a higher initial specialization in routine occupations experienced stronger adoption of information and communication technologies (ICT) and, consequently, a stronger polarization of their earnings and employment structure in later decades. That is, stronger growth in the demand for workers in non-routine cognitive tasks at the expense of workers in routine tasks. This process also affected the relative demand for workers with different educational background as Michaels, Natraj and Reenen (2014) document. The adoption of ICT increased the demand for highly educated workers performing non-routine cognitive tasks at the expense of middle educated workers clustered in routine tasks, leaving the least educated workers largely unaffected.[6]

Our empirical analysis shows that the interplay between two factors are crucial to understand the observed change in the skill-mix of new immigrants: the shift in the relative demand for skills in destinations due to ICT adoption and the shift in the education supply in the origin countries of immigrants.

[5] Autor and Dorn (2013) discuss very briefly the effects of routinisation on the flow of highly skilled *native* workers across local labor markets in the US but do not provide any further evidence on international migrants.

[6] Michaels, Natraj and Reenen (2014) rationale for this is that low educated work in occupations that are both intensive in routine manual and non-routine manual task inputs. The routine-manual task intensive occupations, e.g. jobs in car factories, were largely concentrated in the manufacturing sector. Since routine jobs in traditional manufacturing sectors already declined largely in the 1970s, subsequent ICT growth substituted mainly for middle educated workers in routine cognitive occupations and increased the demand for highly educated workers.

In particular, the change in the education group share in the home countries translates into a similar change in that group's share among immigrants. Thus, education supply predicts well the decrease of the share of low educated immigrants, but explains only a small part of the observed increase in the share of highly educated and considerably over-predicts the increase in the share of middle educated immigrants. This underscores the importance of accounting for the role of relative demand shifts. Accordingly, we estimate a positive effect of routinization on the share of highly educated, a negative effect on the share of middle educated and no effect on the share of low educated immigrants. This result emphasizes the role of routine-biased technological change as a powerful source for demand driven migration. Taken together, supply and demand slightly over-predict the observed change in the share of highly educated workers in Swiss destinations while predicting a small decrease (or stabilization) in the share of middle educated workers.

These results are very robust to a panoply of alternative explanations. In particular, controlling for changes in educational wage differentials, other labor market characteristics and the political environment in the origin countries of immigrants does not affect our estimates for routinization and education supply. They are also robust to controlling for other potential pull drivers, such as ethnic networks, tax incentives or other changes to labor demand, e.g. due to offshoring.

We perform an additional robustness exercise by shedding light on the potential effects of changes in the immigration policy, the remaining factor in the GH framework. To this end, we exploit that immigrants from different origin countries were subject to different immigration restrictions in Switzerland during the three decades from 1980 to 2010. An important differential policy change was in 2002, when the Swiss government implemented a free movement policy for EU workers after 2002 while immigrants from Non-EU countries remained being subject to quotas and a requirement to be highly skilled. This allows to examine whether the skill-mix of new immigrants from EU countries changed differently compared to immigrants from Non-EU countries after the policy was implemented. Then, the estimated effect of the policy can be interpreted as measuring the difference in the effect of a 'free movement policy', as generally prevailing between EU countries and a 'skill-requirement immigration policy', e.g. as adopted by Canada. Our results suggest that the two factors discussed above, routine-biased demand in destinations and education supply in origin countries, are very robust even when we control in very demanding ways for shifts to immigration restrictions. In addition, the estimates for the effect of the policy indicate that the integration of Switzerland into the European labour market after 2002 led, if anything, to a slightly slower change towards highly skilled migrants. That is, the increase in the share of highly educated from the EU was lowered relative to those from other countries. In contrast, the share of low educated immigrants from the EU increased more relative to those from other countries.

Our analysis relates to a larger literature on the selection of immigrants following Borjas (1987)'s adaption of the Roy model.[7] We contribute to this literature, firstly, by documenting that the selection of migrants in OECD countries favoring highly educated over middle educated has accentuated during the past 30 years. Second, based on evidence from local labor markets in Switzerland, we highlight the role of routine-biased technical change as the underlying force shaping these trends. Our findings also extend the literature on job polarization, particularly Autor and Dorn (2013), by showing that the impact of routine-biased technology has important implications beyond the labor market of native workers. In particular, long-run demand trends shape how beneficial the migration to advanced economies is for different types of workers. Immigrants move into (leave) growing (declining) occupations flexibly as do young workers as pointed out by Autor and Dorn (2009). In this sense, our analysis underscores that demand forces, not only short-run

[7] This model has been widely used to highlight the role of differences in returns to skill (Borjas, 2008; Gould and Moav, 2016; Abramitzky, Boustan and Eriksson, 2012) or absolute skill-related earnings differences (Grogger and Hanson, 2011) between host and origin countries to explain selection patterns of immigrants.In addition to Bertoli et al. (2012), only a few other applications use variation over time, highlighting that immigration policies have considerable and rather immediate effects on the number of immigrants (Mayda, 2010; Ortega and Peri, 2014), potentially also affecting the source-country composition (Clark, Hatton and Williamson, 2007).

but also long-run, are of central importance for the migration and location decision of immigrants.

The remainder of this paper is organized as follows. Section 2 motivates a theoretical framework and introduces the empirical strategy. Section 3 discusses the data and stylized facts. Section 4 presents the main results and Section 5 concludes.

5.2 Conceptual Framework and Empirical Approach

5.2.1 Sorting of Immigrants across Local Labour Markets

We consider the stock of migrants from many origin countries in many destinations. The basic ingredient in GH's adaption of the Roy model are separate migration decisions of workers with primary, secondary and tertiary education. Specifically, worker i with education e from origin country o evaluates the utility from migrating to destination j based on the following linear utility function[8]

$$U^e_{i,o,j} = \alpha\left(W^e_{i,j} - C^e_{i,o,j}\right) + \epsilon^e_{i,o,j} \tag{5.1}$$

where $W^e_{i,j}$ and $C^e_{i,o,j}$ are education specific wages and migration costs, respectively, and $\epsilon^e_{i,o,j}$ is an unobserved idiosyncratic term. The wage of worker i is given by

$$W^e_{i,o,j} = \exp\left(\mu_j + \delta^e_j D^e_i\right)$$

where $\exp(\mu_j)$ is the wage of a primary educated worker and $\delta^{e=M}_j$ ($\delta^{e=H}_j$) is the return to secondary (tertiary) education and $D^e_i = 1$ indicates that worker i has education level e. Migrating from origin county o to destination j is a function of fixed costs $f_{o,j}$ and education specific costs $g^e_{o,j}$:

$$C^e_{i,o,j} = f_{o,j} + g^1_{o,j}D^1_i + g^2_{o,j}D^2_i + g^3_{o,j}D^3_i$$

[8] GH also derive predictions using a log-utility function in the fashion of Borjas (1987). In their empirical analysis, however, they show that both linear and log-utility lead to very similar predictions of parameters in the sorting equation, on which we focus here. The reasons is that sorting on log differences in wages is very similar to sorting on level differences in wages in a sample of destinations with similar labour productivity.

GH take Equation (5.1) as a first-order approximation of a more general utility function with $\alpha > 0$ as the marginal utility of income. Staying in the origin country is modeled as the migration costs being zero. Then, assuming (i) that agents base their decision of whether and where to emigrate maximizing their utility and (ii) that errors, $\epsilon^e_{i,o,j}$, follow an i.i.d. extreme-value distribution[9], we can write the log odds of migrating to destination j versus staying in the origin-country o for a worker with education level e as

$$\ln \frac{L^e_{o,j}}{L^e_o} = \alpha \left(W^e_j - W^e_o\right) - \alpha f_{o,j} - \alpha g^e_{o,j} \tag{5.2}$$

where $L^e_{o,j}$ constitutes the population share of workers with education level e from origin country o in destination j and L^e_o is the population share of workers with education e staying in o. Equation (5.2) characterizes scale of immigration, i.e. the number of workers with education e who decide to emigrate to destination j from origin country o. The scale of immigration depends positively on the skill-related wage difference net of migration costs.

The skill composition of immigrants in a destination j from origin country o is just the relative scale of immigration from this country of workers with different educational backgrounds. For concreteness, we can write down separate scale equations for workers with secondary ($e = M$) and tertiary education ($e = H$), take the difference and rearrange to

$$\ln \frac{L^H_{j,o}}{L^M_{j,o}} = \underbrace{\alpha \left(W^H_j - W^M_j\right)}_{\text{(i)}} - \underbrace{\alpha \left(W^H_o - W^L_o\right)}_{\text{(ii)}} + \underbrace{\ln \frac{L^H_o}{L^M_o}}_{\text{(iii)}} - \alpha \underbrace{\left(g^H_{j,o} - g^M_{j,o}\right)}_{\text{(iv)}} \tag{5.3}$$

The sorting equation (5.3) makes predictions about how the number of immigrants with tertiary education relative to those with secondary education

[9] This specification of the disturbance term assumes that *independence of irrelevant alternatives* (IIA) applies among destinations. In our empirical application, we consider different local labour markets within Switzerland as destinations, thus we need only that IIA applies within to the destinations in the sample (Grogger and Hanson, 2011). We can test this assumption by dropping one destination at the time in our regressions and investigating the stability of our estimated coefficients, see Section 5.41.

from a specific origin-country o varies across destinations j. The relative number of highly to middle educated workers in a destination increases, (i) if the wage difference between education groups in the destination j increases, (ii) if the wage difference decreases in the origin-country, (iii) if the supply of highly educated workers increases in the origin country, or (iv) if migration costs fall more for highly educated workers.

5.2.2 Empirical Approach

To characterize the *change* in sorting, we add time indices, t, representing decades, and take first differences

$$\begin{aligned}\Delta\left(\ln\frac{L^H}{L^M}\right)_{j,o,t} &= \alpha\Delta\left(w^H - w^M\right)_{j,t} - \alpha\Delta\left(w^H - w^M\right)_{o,t} \\ &\quad + \Delta\left(\ln\frac{L^H}{L^M}\right)_{o,t} - \Delta\left(g^H - g^M\right)_{j,o,t}\end{aligned} \tag{5.4}$$

where Δ represents differences over decades, t, $\Delta x_t = x_{t+1} - x_t$. To take this expression to the data, we need information on wages in Swiss commuting zones, the destinations in our case, and origin-countries as well as information on the skill supply in origin countries and relative migration costs. However, using wages in destinations is problematic, as wages may be endogenous to immigration.[10]

We suggest a different route and use a proxy for local, relative demand shifts that is unaffected by immigration and which also allows testing the role of routine-biased technical change directly. Autor and Dorn (2013) show that a region's 'initial' share of workers employed in routine-intensive occupations, denoted by $RSH_{j,t}$, is a good indicator for subsequent relative demand shifts, affecting educational wage differentials and inducing a polarization of the local employment structure.[11] The intuition is as follows. Comput-

[10] The existing literature studying drivers and the selection and sorting of immigration largely ignored this concern. One notable exception is Mayda (2010) who uses lagged income measures to investigate the scale of immigration.

[11] The idea of routine intensity as a proxy for relative demand shifts affecting the wage differential of workers with different educational backgrounds and skills has found wide application in the literature on skill-biased technical change and job polarization. See

ers (or ICT more generally) are a close substitute for workers employed in jobs with a large share of routine manual or routine cognitive tasks, such as assembly line workers or bank clerks who, typically, had middle earnings. The advent of computers after 1980 and their continuously falling price over the past decades has lead firms to substitute computers for these workers and has driven down their relative wages. On the other hand, ICT complements workers employed in managerial or professional occupations, typically highly paid and working on non-routine abstract tasks. The adoption of computers increased the demand for these workers, raising their wages and employment. Indeed, Autor and Dorn (2013) show for the U.S. that regions with a larger, initial share of employment in routine occupations in 1980, experienced stronger wage and employment polarization subsequently.[12]

Michaels, Natraj and Reenen (2014) find that this process is also reflected in the changing demand for workers with a different educational background; ICT replaced workers with a middle education level, which were mainly employed in routine occupations, while it increased the demand for highly educated workers employed in non-routine cognitive occupations. On the other hand, since many low educated workers are employed in non-routine manual tasks they have been much less affected by ICT adoption.[13]

Consequently, we expect that regions with a larger initial share of routine employment experience larger positive demand shifts for highly educated workers relative to middle educated workers inducing their wage difference

Acemoglu and Autor (2011) for an overview of the relevant literature. After Autor, Levy and Murnane (2003)'s seminal contribution showing wage and employment trends of workers with different routine-task content in their jobs for the U.S., similar trends were documented for the U.K. (Goos and Manning, 2007), Europe (Goos, Manning and Salomons, 2014) and Germany (Dustmann, Ludsteck and Schönberg, 2009; Spitz-Oener, 2006) showing some connection to the routine intensity. We explain in Section 3 in detail, how we measure $RSH_{j,t}$.

[12] In their framework, Autor and Dorn (2013) assume that capital is fixed across regions, while labour is mobile. The larger increase in wage inequality in more routine intensive regions induces highly educated workers to move to those regions. Higher inflow of workers, in turn, drives up local prize levels and slows the increase of nominal wages of highly educated in those areas. Labor flows slow down, as real wages equalize across regions.

[13] In Section 3, we corroborate this link between education levels and the task content of their occupations in the Swiss case.

to increase. Thus, we can write down the empirical versions of the sorting equation (5.4) for highly educated relative to middle educated workers where we substitute changes to wage differences in destinations with a region's routine share:

$$\begin{aligned}\Delta\left(\ln\frac{L^H}{L^M}\right)_{j,o,t} = \; & \alpha RSH_{j,t} - \alpha\Delta\left(w^H - w^M\right)_{o,t} \\ & - \alpha\Delta\left(g^H_{j,o} - g^M_{j,o}\right) + \alpha\Delta\left(\ln\frac{L^H}{L^M}\right)_{o,t}\end{aligned} \tag{5.5}$$

While we also estimate Equation (5.5) directly, for the main part of our analysis we take the share of a skill group as the dependent variable, which allows for an more intuitive interpretation of the estimated coefficients[14]:

$$\begin{aligned}\Delta EDUSH^e_{j,o,t} = \; & \beta^e_1 RSH_{j,t} + \beta^e_2 \Delta EDUSH^e_{o,t} + \beta^e_3 \Delta X^e_{o,t} \\ & + \alpha_c + \alpha_o + \alpha_t + \epsilon_{j,o,t}\end{aligned} \tag{5.6}$$

where $\Delta EDUSH^e_{j,o,t}$ is the decennial change in the education group share of new immigrants from origin-country o in destination j. We estimate separate regressions for each education group $e \in \{L, M, H\}$, i.e. low, middle and highly educated, and pool multiple decades as stacked first-differences. $\Delta EDUSH^e_{o,t}$ is the change in education group shares in origin country population of immigrants and $\Delta X^e_{o,t}$ represents proxies for changes to relative wages or earnings inequality and a set of additional controls varying by decade and origin country. Fixed effects on the level of cantons, α_c, absorb linear trends in other pull forces for destinations in the same canton, e.g. differences in institutional arrangements. Origin country fixed effects, α_o, absorb linear trends in other push forces and decade fixed effects, α_t, absorb differences across decades. If the routine share of a location captures a relative demand shift for highly relative to middle educated workers, we expect that $\beta^H_1 > 0$ and $\beta^M_1 < 0$ while the sign for β^L_1 is ambiguous according to the reasoning

[14] We will provide robustness checks for the choice of the functional form in Section 5.41. We also discuss the importance of the linear utility assumption and potential violation of the IIA assumption.

above. Furthermore, we expect that β_2^e is positive and close to one for all education groups.

In our empirical analysis, we augment Equation (5.6) in various ways. Although most scholars point out that technology induced routine-biased technical change was a dominant force characterizing the labor market of OECD countries (Goos, Manning and Salomons, 2014), trends in the skill-mix of immigrants could also have been induced by other forces. Therefore, we will test different measures of offshorability of a destination's employment, for instance as suggested by Goos, Manning and Salomons (2014). Additionally, we will investigate the role of immigration policy by incorporating indicators for the differential degree of openness of the Swiss labor market for different origin countries and different decades (see Section 5.43).

5.3 Data, Measurement and Stylized Facts

5.3.1 Data Sources and Definitions

Recent Immigrants

We measure the educational composition of new immigrants in Swiss local labor markets using data from the Census 1980, 1990 and 2000 in combination with the pooled Structural Survey 2010 to 2012 provided by the Federal Statistical Office (FSO).[15] We classify individuals as *new immigrants*, if they were born abroad and arrived less than 5 years before the Census wave.[16] Among new immigrants, we can distinguish 30 different *origin countries* based on the country of residence 5 years ago.[17]

[15] In contrast to the Census, which constitutes a compete count of the population, the Structural Survey represents a 3% sample of the total resident population (Swiss Federal Statistical Office, 2011). We pooled the structural surveys from 2010 to 2012 to gain more accuracy.

[16] For 2010 to 2012, the information on the year of arrival is missing in some entries. In this case, we classified foreign-born residents as recent immigrants if they had a short-term residency permit (B, L).

[17] Using the last residency country reflects more closely the immigration decision in the sense of GH compared just using the country of birth as origin. However, the correlation between the two classification of origin is very high in our data. As the Census does not distinguish different places of origin for immigrants from Ex-Yugoslavia and the former Czechoslovakia, we aggregate immigrants from all available countries of former

Individuals are classified into three education groups using the International Standard Classification of Education (ISCED). *Highly educated* individuals hold a tertiary degree (ISCED 5 and 6), whereas *middle educated* individuals hold a degree from a secondary school (ISCED 3 and 4). *Low educated* individuals are those with compulsory education only or less (ISCED 0, 1 and 2).

As *destinations* we use the 106 Swiss commuting zones (CZs) as defined by the FSO (Schuler, Dessemontet and Joye, 2005). These CZs are a good approximation of local labor markets and are constructed such that the majority of people commute to work within its boundaries.

Our base sample consists of new immigrants who are older than 15 years and have non-missing information for education and their place of living. When we show stylized facts for workers, we focus on workers between 15 and 65 years and measure labour supply in full time equivalents based on weekly hours worked. Observations from the Structural Survey are weighted using the official sampling weights. Using these definitions, we collapse our dataset into year, CZ, origin country and education group cells of new immigrants. Technically, this would leave us with a sample of 9540 (3 decades $\times$ 30 origin countries $\times$ 106 CZs) observations for each education group when first differencing the education group shares across decades as in our baseline specification. One issue is the presence of zero or missing bilateral migration stocks. As GH point out, based on the law of large numbers, theory would predict all bilateral stocks to be positive, though some might be very small. Zero migration stocks may occur in finite populations if bilateral migration probabilities are very small. We deal with this by setting all empty cells to zero for 1980 to 2000, where we have a full population inventory. Since the structural surveys are not full inventory counts, we treat empty cells for 2010 as missing. Thus, our final sample includes 5304 observations.[18]

Yugoslavia and aggregate immigrants from the Czech Republic and Slovakia in the Structural Surveys 2010 to 2012. Note that we focus on new immigrants, rather than the total of foreign born, as they represent best the "marginal international migrant" moving from an origin country to a destination. Thus, they are most affected by the forces that we discuss in this paper.

[18] If all education groups in a CZ-origin country pair are missing, the calculation of education shares is mathematically not defined and, hence, such a CZ was treated as a missing

Origin Country Information

We use various sources to control for origin country push drivers. To calculate *education group shares* in origin countries, we use the data from Barro and Lee (2013) which reports the percentage of the population in different levels of educational attainment. We define ‘no schooling attainment’ and ’primary schooling attainment‘ as *low educated*, ‘secondary schooling attainment’ as *middle educated* and ’tertiary schooling attainment‘ schooling attainment as *highly educated*.[19]

To proxy for wage differences between education groups we use Gini coefficients from the UNU-WIDER World Income Inequality Database (WIDER), Version 2.0c (May 2008). Generally, we only included inequality measures based on disposable income. As there were numerous sources for Gini coefficients for some countries, we computed averages for a given country-year cell.[20] As an alternative, we also use income based Gini coefficients from the Luxembourg Income Study (LIS) ‘Key Figures’ (Version 3), which are available for a subset of decades and only for developed countries.[21] For some additional robustness checks, in which we directly estimate the sorting equation (5.5), we constructed education specific wage measures using percentile ratios from the LIS.

As additional controls, we use per capita GDP from Heston, Summers and Aten (2011) and a country’s polity IV score as a measure for civil rights protection and democratization from Marshall, Gurr and Jaggers (2014). To measure exposure to clivil conflict, we retrieve information whether a country

observation. In additional results in Table 5.A.6 in the appendix, we demonstrate that our findings are robust to dropping these empty migration cells.

[19] We use the population weighted means from Croatia, Serbia and Slovenia to calculate the education measure for ’Ex-Yugoslavia‘ and of the Czech Republic and Slovakia for measure of ‘Czechoslovakia’.

[20] The averaging resulted usually in little differences compared to relying only on a single datasource. For other countries such as China or India, inequality measures based on disposable income were scant or unavailable. In these cases we also included studies that calculated Gini coefficients based on consumption or net income data.

[21] Data for the majority of origin countries in the Non-EU group is missing in the LIS data. See Appendix 3 for more details.

was exposed to a conflict in a decade with at least 25 battle-related fatalities from UCDP/PRIO's "Armed Conflict Dataset" (UCDP/PRIO, 2015).

Measuring Routine Intensity of Swiss Commuting Zones

We measure a CZs specialization in routine tasks by its occupational composition of employment. To this end, we merge job task requirements from the US Department of Labor, Employment and Training Administration (1977)'s Dictionary of Occupational Titles (DOT) to occupations available in the Swiss Censuses.[22] Following Autor and Dorn (2013) we combine the task measures from the DOT to create an indicator of routine-intensity by occupation:

$$RTI_k = \ln\left(T^R_{k,1980}\right) - \ln\left(T^M_{k,1980}\right) - \ln\left(T^A_{k,1980}\right) \tag{5.7}$$

where $T^R_{k,1980}$, $T^M_{k,1980}$ and $T^A_{k,1980}$ are the routine, manual and abstract task inputs in occupation k in 1980.[23] We then calculate for each CZ the employment share in routine-intensive occupations, $RSH_{j,t}$, subsequently referred to as routine share:

$$RSH_{j,t} = \left(\sum_{k=1}^{K} L_{jkt} \times 1[RTI_k > RTI^{P66}]\right)\left(\sum_{k=1}^{K} L_{jkt}\right)^{-1} \tag{5.8}$$

where L_{jkt} is the employment in occupation k in commuting zone j and decade t. $1[.]$ takes a value of one if occupation k is in the top employment-weighted third of routine task-intensity.

In our baseline regressions, we follow the instrumental variable strategy of Autor and Dorn (2013) to isolate the long-run component of a commuting zone's current routine share. That is, we use the industrial composition

[22] Autor and Dorn (2013) provide a measure for routine, abstract and manual task content for US 2000 census occupations (occ2000) from the DOT 1977. These three task aggregates were collapsed from originally five task measures first used in Autor, Levy and Murnane (2003). We use a crosswalk from the US National Crosswalk Service Center to match these variables to the International Standard Classification of Occupations (ISCO-88) available in the Swiss Census.

[23] Each task is measured on a one to ten scale, with ten meaning that the task is most heavily used in this occupation.

of CZs in the first available Census (1970) as instruments for the observed routine share in later decades:

$$\widetilde{RSH}_j = \sum_i \frac{L_{i,j,1970}}{L_{j,1970}} \times RSH_{i,-j,1970}. \tag{5.9}$$

$\frac{L_{i,j,1970}}{L_{j,1970}}$ represents employment in industry i as a share on total employment in CZ j in 1970. $RSH_{i,-j,1970}$ is the routine share of employment in industry i in all CZs except j in 1970.[24] Column 1 in Table 5.1 reports the first stage estimates, regressing the stacked routine shares 1980, 1990 and 2010 on the instrument, fixed effects (by origin country, year and canton) and some baseline controls.[25] Column 2 to 4 show estimates by decade. The declining magnitude of the coefficients illustrates how the predictive power of initial routine intensity in 1970 declines over time which has also been noted by Autor and Dorn (2013).

5.3.2 Stylized Facts

Trends in the Skill-Mix and the Selection of Immigrants. Figure 5.1 shows the evolution of the three education group shares (Panel A to C) among new immigrants in Switzerland (blue solid line) and the average shares in the home country populations of Swiss immigrants (red dashed line) between 1980 and 2010. As in the case of OECD destination countries (cf. Figure 5.1), we see that highly educated individuals were over-represented among immigrants (16%) compared to non-emigrants (6%) in 1980. Middle educated individuals were under-represented (22% vs. 32%) and low

[24] We calculate the routine share using total employment in 1970 and later decades. If established immigrants were significantly clustered in routine occupations and new immigrants relied largely on existing networks of compatriots in their location decision, we would observe a spurious correlation between the routine share and subsequent immigration inflow. To address this concern, we will show robustness checks including a CZ's population share of immigrants from an origin country in 1970.

[25] This regression also includes the change in the share of highly educated in the origin countries of immigrants, see specification in Column 1 of Table 5.1. Unsurprisingly, the first stage results are very similar irrespective of the independent variable in the second stage. The first stage results presented here are for a regression with the change of the share of highly educated new immigrants as independent variable in the second stage.

Table 5.1: First Stage Estimates of Baseline Regression (in Table 3, Panel A, Column 1)

Sample	1980-2010	1980-1990	1990-2000	2000-2010
$\widetilde{RSH}_{j,1970}$	0.788	1.172	0.644	0.536
	(0.088)***	(0.121)***	(0.079)***	(0.052)***
R^2	0.527	0.720	0.618	0.614
Observations	4543	1964	2079	500
F	80.362	93.569	66.583	104.896

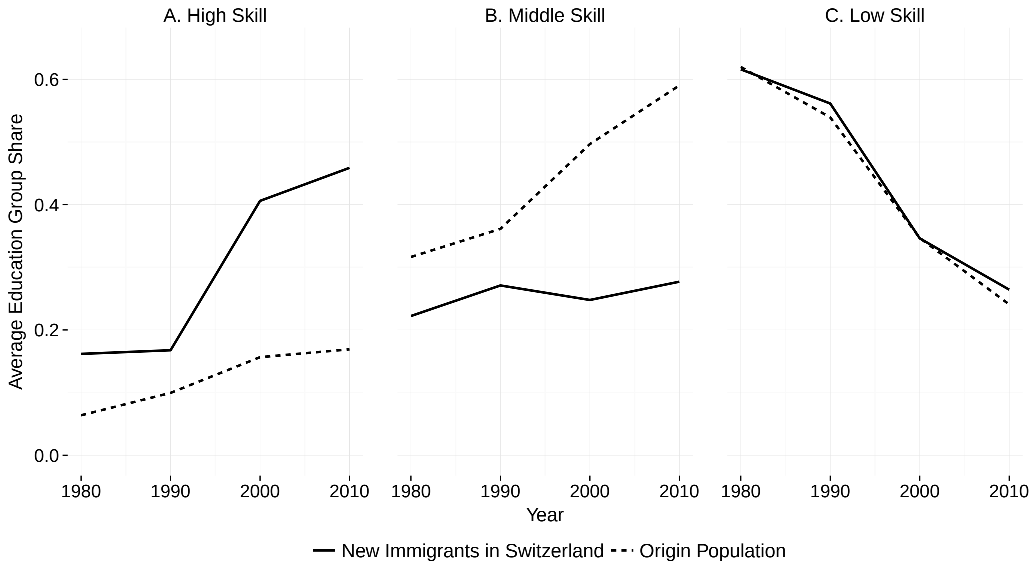
Note: ***, **, *, denote statistical significance at the 1%, 5% and 10% level. Robust standard errors (clustered by CZ and origin country) are given in parentheses. Regressions include fixed effects for Cantons, origin countries and are weighted using the total number of recent immigrants from origin country o in destination j at the beginning of the decade. Column 2 to 4 show separate estimates for every decade.

educated about equally represented. By 2010, the share of highly educated immigrants increased strongly to 46% while the share of middle educated immigrants increased much slower to 28%. In contrast, in the origin countries the population weight of middle educated increased much more than that of highly educated. Finally, the share of low educated among immigrants decreased almost in parallel to the corresponding share in the home country population.

Figure 5.2 breaks the same information down by different origin countries: it compares the average education group shares of new immigrants with the education group shares in their home countries in 1980 (panel A) and in 2010 (panel B).[26] In addition, each panel plots the mean of all countries for that year, using the numbers of immigrants by origin country as weights. This shows, first, that the trends documented above are not driven by a small number of origin countries. Immigrants from most source countries were positively selected in 1980, showing higher shares of highly educated and lower shares of middle and low educated in destinations compared to home

[26] For completeness, Table 5.A.2 in the appendix presents the full list of origin countries, their share of new immigrants in 1980 and 2010, the education group shares among new immigrants and in the home country population and the changes in the education composition between 1980 and 2010.

Figure 5.1: Skill Group Shares among New Immigrants in Switzerland and among the Population in their Origin Countries (Averaged), 1980 - 2010



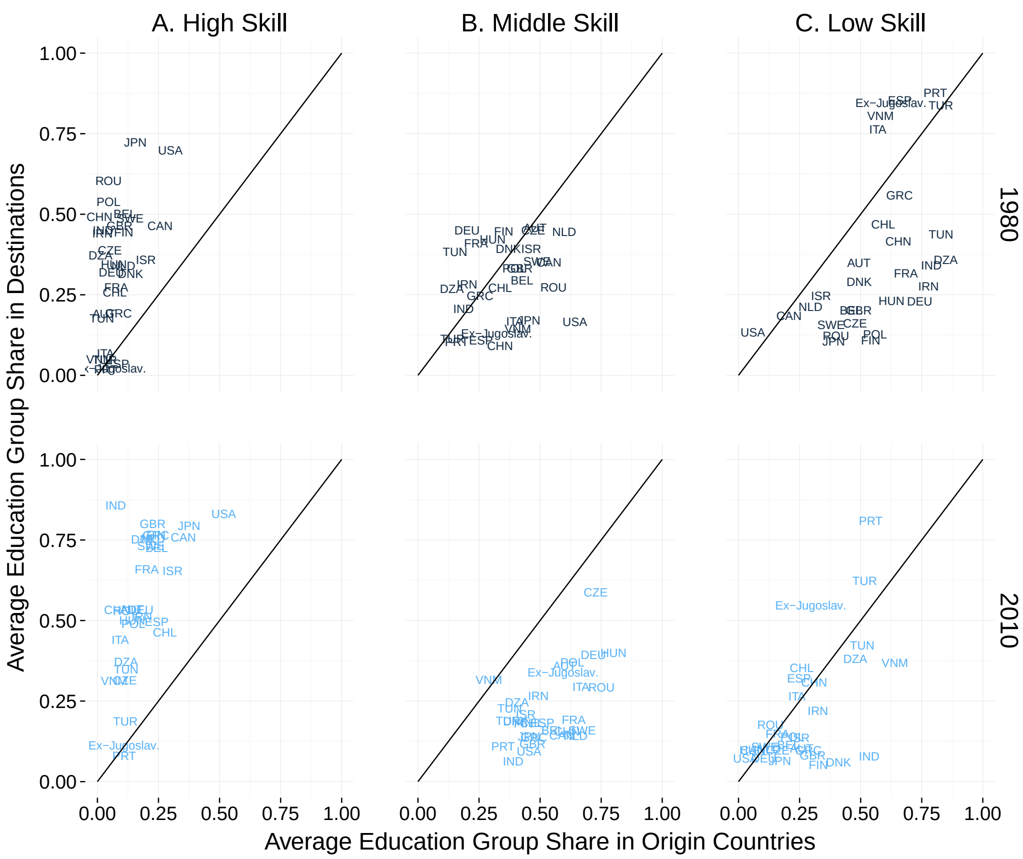
Note: The education groups shares for new immigrants are computed using Swiss Census Data 1980-2010. The education group shares in the origin countries are computed using the population weighted average across origin countries from Barro and Lee (2013) with the number of immigrants from a country-year in the Swiss Census as weights.

countries. Only from a few countries were immigrants negatively selected (Spain, Italy, countries from Ex-Jugoslavia, Portugal).

Second, the comparison with Panel B shows *how selection changed* between 1980 and 2010. If educational upgrading in the source countries translated one to one to individuals migrating to Switzerland, we would see the scatter cloud moving parallel to the 45-degree line; towards the upper right corner in the case of the highly and middle educated and to the lower left corner for the low educated group. However, this happened only for the low educated group: Its share declined significantly in each origin country and also among immigrants with the mean share of low educated slightly over-represented among new immigrants in 2010.

In contrast, the Figure shows remarkably clearly how the positive selection of highly over middle educated became even more accentuated between 1980 and 2010. For most origin countries, the share of highly educated in-

Figure 5.2: Skill Group Shares among New Immigrants in Switzerland and among the Populations in their Origin Countries, 1980 and 2010



Note: The education groups shares for new immigrants are computed using Swiss Census Data 1980-2010. The education group shares in the origin countries are taken from Barro and Lee (2013). "Mean1980" and "Mean2010" represents the average weighted education groups shares using the number of immigrants from an origin country in a 1980 and 2010, respectively, as weights.

creased more among immigrants than in their home country population.[27] The share of middle educated, in comparison, moved away from the 45-degree line to the right, meaning that middle educated immigrants became even more underrepresented in Switzerland.

The fact that these trends concern immigrants from many different origin countries is astonishing because those countries were subject to different la-

[27] For instance, this share increased by staggering 57 percentage points among new immigrants from Greece. As the share of highly educated was low among Greek immigrants in 1980, this raises the issue of potential mean reversion in destinations, we will carefully demonstrate the robustness of our results to initial conditions.

bor market environments and institutional settings including migration policies. This suggests that something more fundamental is at play affecting immigrants from all origin countries.

Routine intensity of occupations by education level Next, we analyze the occupational task content of workers in 1980 in Switzerland. As Michaels, Natraj and Reenen (2014) point out, the degree to which workers with different education levels are clustered in routine (non-routine) intensive occupations sheds light on the substitutability (complementarity) of workers with different education levels with ICT technology.

Table 5.2: Task Content of Education Groups

	Task Content			
Skill Group	abstract	routine	manual	RTI
	A. Total Workforce in 1980			
High	1.128	-0.299	-0.195	-0.368
Middle	-0.056	0.123	-0.095	0.117
Low	-0.328	-0.062	0.202	-0.030
	B. Recent Immigrants in 1980			
High	1.082	-0.179	-0.213	-0.318
Middle	-0.315	0.163	0.190	-0.030
Low	-0.561	0.221	0.708	-0.155

Note: Task measures taken from DOT as described in Section 3. Routine intensity (RTI) calculated as in Equation (5.7). Task measures and RTI scores are first standardized to have mean zero and standard deviation one in the entire workforce. Then, averages are computed over all workers in an education group using employment weights. Agricultural workers have been omitted from this table. Swiss Census 1980.

Table 5.2 (Panel A) shows the average intensity in abstract, routine, manual tasks and the RTI compound-measure explained above, using the distribution of total employment in each education group across occupations in 1980. The first row in Panel A shows that tertiary educated workers were largely employed in occupations with the highest abstract task content and the lowest routine and manual task content. Workers with a middle education were mostly employed in occupations with a high routine task content and low task intensities for both manual and abstract tasks. Finally, the picture for low educated workers is more mixed: They have very low scores in abstract tasks but are only slightly below average in routine tasks and have very high average manual task content. As Michaels, Natraj and Reenen (2014) point out, this illustrates that the adoption of ICT capital might have no clear effects on the least educated group while increasing the demand for highly educated workers at the expense of middle educated workers.

Immigrants might not cluster in exactly the same occupations as natives with respect to the education level. Thus, routine-biased demand shifts could affect them differently. To analyze this, Panel B of Table 5.2 shows the task content of occupations for new immigrants in 1980. While highly educated immigrants worked in occupations with similar tasks as characterized for the total workforce, middle and low educated immigrants worked in occupations with both above average routine and manual task inputs. The importance of routine tasks relative to manual tasks was highest among middle educated immigrants whereas low educated workers had a very high manual task intensity. Thus, we would suspect that routinization should mostly increase the demand for highly educated immigrants. On the other hand, routinization might lower the demand for both middle and low educated groups, with smaller effects on the low educated who are shielded in more manual intensive tasks.

Job and wage polarisation and skill trends in destinations. Although the polarisation of employment and the paralleling change in wage-inequality has been documented extensively for most developed economies (e.g. Acemoglu and Autor, 2011; Goos, Manning and Salomons, 2014), Dustmann, Ludsteck and Schönberg (2009) point out that there are potentially important cross-country differences in how the adoption of ICT affects the occupational employment and wage structure. For Switzerland these trends have not been documented satisfactorily for our purpose.[28] We summarize the most important facts in what follows.

Figure 5.3 shows the change in the employment shares of ISCO main occupation groups between 1980 and 2010, separately for natives and new immigrants.[29] Occupations are ranked by their mean log hourly wage using pooled data from the first two available waves of the Swiss Labor Force Survey (SLFS) 1991 and 1992. Employment has polarized considerably in the case of natives and even more so for recent immigrants. For example, the fraction of new immigrants employed in abstract occupations like managers (ISCO 1) increased by roughly 12 percentage points, from 2.7% in 1980 to almost 15% in 2010, whereas for native workers it grew by 5 percentage points (from 6% to to 11%). On the other hand, employment in routine-manual occupations such as craftsmen (ISCO 7) fell sharply for new immigrants, from over 41% in 1980 to 15% in 2010. For natives, it changed more modestly from 24% to about 14%. The fraction of workers in service and elementary occupations (ISCO 5 and 9) stayed roughly the same for both new immigrants and natives. Summing up, job polarisation seems to be more pronounced among new immigrants compared to natives.[30]

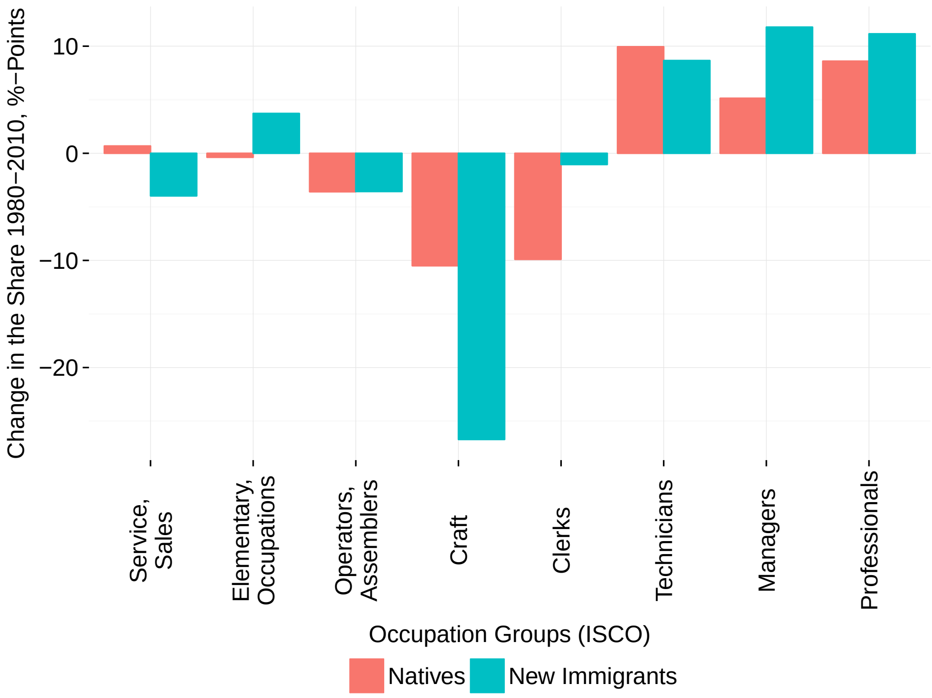
This employment polarization has been accompanied by wage polarization. Figure 5.4 shows the change in mean log hourly wages for ISCO main occupation groups between 1991 and 2010 using SLFS data. As the sample size is considerably smaller in this survey, we aggregate ISCO main occu-

[28] Oesch and Menés (2011) compare job polarisation in Switzerland, the UK, Spain and Germany. For Switzerland, the rely on a relatively small sample from the Swiss Labour Force Survey and a relatively short time span between 1991 and 2008, i.e. when computerization was already well underway. Splitting the employment distribution into earnings quintiles, they find employment growth only at the top of the earnings distribution. Müller, Asensio and Graf (2013) find an increase in wage-inequality at the top of the wage distribution relative to the middle but do not rely on occupations for their analysis as we do here. Consequently, they do find very different results for wage changes at the bottom.

[29] For completeness, Table 5.A.1 in the appendix shows the corresponding employment shares for each group and for the total workforce in all decades 1980-2010 and the change 1980 to 2010.

[30] This results complement the findings of Autor and Dorn (2009) who show that entry into abstract occupations is most pronounced among young workers in the US. This shows that immigrants are an important channel of adjustment (next to the entry of new cohorts) to long-run changes in skill-demand.

Figure 5.3: Change in Employment Shares of Occupation Groups, 1980 - 2010, Natives and New Immigrants

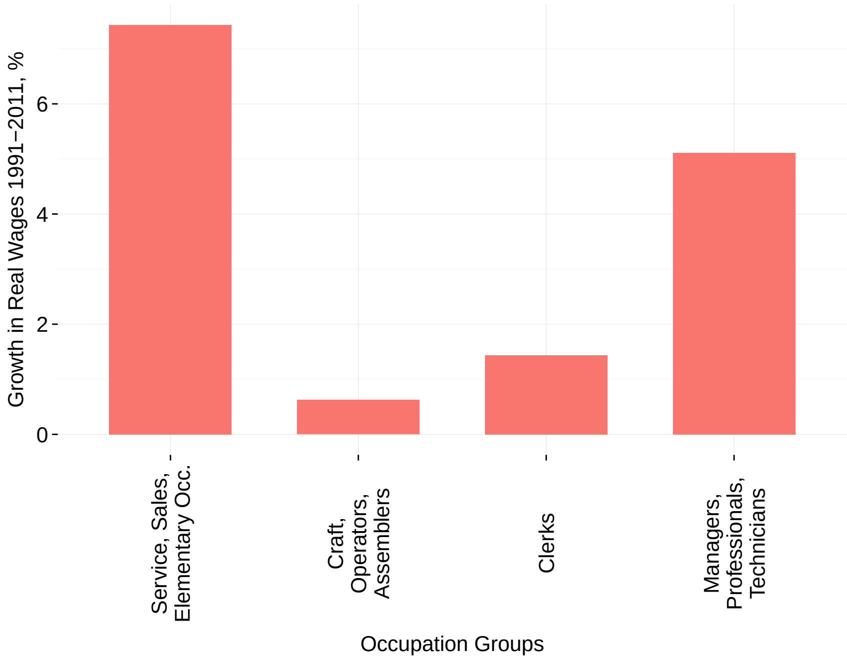


Note: The graph shows for each ISCO main occupation group (omitting agriculture) the average decennial growth of its employment share (in full time equivalents) on the total labour force between 1980 and 2010. Occupation groups are ranked by their median wage taken from the Swiss Labor Force Survey (1991 to 1992 pooled). Employment data from Swiss Census 1980 to 2010.

pation groups into four groups according to their task content.[31] Evidently, wage growth is most pronounced in low paying occupations and at the top with more moderate gains in both routine-manual (craft, operators, assemblers) and routine cognitive occupations (clerks). As highly educated workers are mostly clustered in abstract occupations and middle educated in routine occupations, we would expect that the wages earned by highly educated workers increased relative to those with a middle education.

[31] These are non-routine occupations (service, sales and elementary occupations), routine manual occupations (craft workers, machine operators and assemblers), routine cognitive occupations (clerks) and non-routine abstract occupations (managers, professionals and technicians). We ranked those groups again by their mean log wages in 1991.

Figure 5.4: Change in Real Mean Wages of Occupation Groups, 1991 - 2011



Note: Change in mean log hourly wages by broad occupation groups between 1991 and 2011. Service occupations is the aggregate of ISCO 5 and ISCO 9 occupations, Craft/Operators/Assemblers is the aggregate of ISCO 7 and ISCO 8 occupations, Clerks is ISCO group 4 and Managers/Profs/Techns is the aggregate of ISCO groups 1 to 3. The years 1991-1992 and 2010-2011 are pooled. Swiss Labour Force Survey data.

These national trends have a clear geographical echo as posited and documented for the US by Autor and Dorn (2013); regions with larger initial employment in routine occupations in 1980 experience more adoption of ICT capital and, consequently, larger wage and job polarisation. In turn, the increase in the earnings of highly educated workers relative to middle educated induces larger (smaller) immigration of highly (middle) educated workers. Figure 5.5 shows simple correlations between the change in education group shares among new immigrants in CZs between 1980 and 2010 and the routine employment shares of CZs in 1980.[32] Indeed, the share of highly educated

[32] Specifically, we run the following OLS regressions, separately by education group: $\Delta EDUSH^{e}_{j,1980-2010} = \alpha^e + \beta^e RSH_{j,1980} + \epsilon_j$. Regressions are weighted with the number of new immigrants in a CZ in 1980. The estimated coefficient of the routine share for each education group is given below the scatter plot. Note that the dependent variable is constructed from the full sample of new immigrants, not only workers.

immigrants increased much more in areas with a larger routine employment share in 1980. In contrast, the share of middle educated increased more slowly or even decreased in these areas.[33] In Zurich, for instance, the share of tertiary educated new immigrants increased from 22% in 1980 to a staggering 70% in 2010, whereas the share of middle educated declined from 24% to 17%. In Lucerne, in contrast, which had a smaller routine employment share in 1980, the share of highly educated increased more modestly relative to Zurich (from 11% to 39%), whereas the share of middle educated even increased (from 20% to 35%). For low educated immigrants, the overall correlation is also positive, but the plot looks more noisy than for the other two groups. Although these correlations are supportive for our predictions, there might be many unobserved factors which could also drive these results. The following section analyze these factors more rigorously.

5.4 Results

5.4.1 Baseline Factors of the Skill-Mix of New Immigrants

This section takes the empirical counterpart of the sorting equation, Equation (5.6), to the data. To do so, we pool multiple decades as stacked first-differences and use two-stage least squares, with $RSH_{j,t}$ instrumented by a CZ's historical industry specialization in routine intensive occupations as specified in Equation (5.9).[34] Table 5.1 reports the results when the dependent variable is, alternatively, the share of the highly (panel A), middle (panel B) and low (panel C) educated new immigrants. All regressions are weighted with the number of immigrants from origin country o in destination

Results for new immigrant workers, however, are very similar. Differences between the full population of new immigrant and its workers are discussed in Table 5.2.

[33] We find similar results, if we regress instead the change in selection in a destination between 1980 and 2010 on the routine share in 1980. The change in selection, in this case, is the difference in the actual observed change of an education group share minus the average change in the same education group share in the origin countries of immigrants.

[34] The F-Statistic in case of all 2SLS estimates are well above the conventional threshold of 10 which indicates that our IV strategy works generally well (Stock, Wright and Yogo, 2002). Table 5.A.3 in the appendix presents additional estimates with similar specifications as in table 5.1 using OLS.

Figure 5.5: Change in Skill Group Shares of New Immigrant Workers in CZs 1980-2010 and Routine Employment Share of CZs in 1980

Note: Scatterplot of average change in the education groups shares of new immigrant between 1980 and 2010 in a CZ against share of routine employment in CZ in 1980. The size of the circle reflects the total number of immigrants in a CZ in 1980. The solid (dashed) line represents the predicted change in the education group share (5% confidence interval) from the following OLS regression model for each education level e: $\Delta EDUSH^e_{j,1980-2010} = \alpha^e + \beta^e RSH_{j,1980} + \epsilon_j$. An estimate of β^e and its standard error is shown below each plot.

j at the beginning of the decade. Standard errors are two-way clustered by CZ and origin country (Cameron, Gelbach and Miller, 2011).[35]

Column 1 shows the results from the regression that includes only the routine share, $RSH_{j,t}$, the change in the education structure in the source countries, $\Delta EDUSH^e_{o,t}$ and fixed effects for origin countries, cantons and decades. The coefficient of $RSH_{j,t}$ shows that destinations with a larger initial share of employment in routine occupations experience larger employment growth of highly educated immigrants. The opposite is true for the share of middle educated, while leaving the least educated group unaffected as in Michaels, Natraj and Reenen (2014). More specifically, a region with a one percentage point higher routine share in 1980 would subsequently experience a 0.34 (-0.3) percentage point higher (lower) increase in the share of highly (middle) educated immigrants in each decade between 1980 and 2010. The coefficients of education supply, $\Delta EDUSH^e_{o,t}$, differ to some de-

[35] There are only 26 Cantons which is why we chose to cluster on the level of CZs (and origin countries). Results are similar when clustering at the level of Cantons.

gree across education groups, but the confidence intervals of the estimates for both high and low educated includes unity whereas the estimate of middle educated is only slightly but significantly below this number. That is, a one percentage point change in the share of an education group in the origin countries leads to almost the same change in the destinations as posited in GH's model (see e.g. Equation (5.4)).

Column 2 adds the change in the Gini index from UNU-WIDER and the growth of GDP per capita. This is a first way to control for changes to returns to skill and general income levels in the origin countries. Both coefficients have the expected sign. An increase in the returns to skill reduces the share of highly educated immigrants and increases the share of low educated. On the other hand, higher growth in the source countries seems to matter mostly for low educated by reducing their share. The coefficients for the relative demand effect and the supply effect are statistically not different from those in Column 1.[36]

In Column 3, we test the robustness of our preferred measure for the returns to skill in the source countries by using an alternative Gini measure from the LIS. This leaves us with a considerably smaller sample, dropping most developing countries.[37] Importantly, this has no substantial effects on the estimates of the routine share and the change in the education group shares in the source countries. The estimates for GDP per capita and the Gini index are larger than in Column 2 but keep the same sign.

In Column 4, we use again the UNU-WIDER Gini index and include two other controls which are available for most origin country-year observations. First, we include changes in the Polity IV score across decades to control for the political conditions and civil liberties in the origin countries of immigrants which could affect skill groups differentially (Grogger and Hanson, 2015).

[36] One concern its that the decennial changes in control variables from an origin country could be correlated with the outflow of emigrants from this country to Switzerland. Taking the control variable in levels (at the beginning of the decade) rather than decennial change, as reported in Table 5.A.4 of the online appendix, does not change the estimates of the routine share and the educational supply measure.

[37] We only have LIS Gini measure for European countries (most Eastern European countries only after 1990) and for Canada, Israel and the US. For Ex-Jugoslavia, the Gini index in 1990 is approximated with the corresponding number from Slovenia.

Second, we include a control for whether an origin country has been affected by civil conflict, which could lead to large refugee waves. Particularly the Balkan wars could be an important push driver affecting the skill-mix of an important share of new immigrants during the 1990s. Indeed, while the estimate of the Polity IV score is not significant, the share of low educated increased considerably more (15 percentage points) among new immigrants from conflict affected areas lowering the share in the highly educated group.[38]

Lastly, Column 5 includes interaction terms between origin country and time fixed effects in order test the robustness of our results to other omitted push-factors and potential deficiencies in our measures of skill returns. Importantly, these fixed effects also absorb all potential variation from immigration policies, as those varied only over origin countries and decades. We discuss this point more extensively in Section 5.43. Reassuringly, our estimates for the $RSH_{j,t}$ coefficient prove to be robust even to this demanding specification.

Overall, a consistent picture emerges from these findings. The effects of the educational composition in the origin countries and the routine-biased demand shifter are very similar across different specifications and samples. Together, they explain well the large increase in the share of highly educated immigrants at the expense of middle educated immigrants as we show next. On the other hand, changes to the labor market conditions in the origin countries, including relative wage changes, may matter but play a secondary role for the differential migration decision of workers with different educational backgrounds. This has also been shown for the general magnitude of immigration flows (see e.g. Mayda, 2010).

[38] We also tested the inclusion of the change in the unemployment rate, as an additional control for origin country labor market characteristics which is only available for a subset of countries. As many authors, e.g. Belot and Hatton (2012), point at the importance of credit constraints, influencing the immigration decision of poor households, we also checked the influence of including the change in the inverse of GDP per capita squared, a proxy for poverty suggested by Clark, Hatton and Williamson (2007). Neither of these variables was significant or had a substantial influence on the coefficients of the other variables shown above.

Table 5.1: Determinants of the Change in Skill Group Shares, 1980 - 2010, 2SLS

	(1)	(2)	(3)	(4)	(5)
	A. Dependent variable: Change in Share of Highly Educated				
$RSH_{j,t}$	0.340	0.362	0.412	0.362	0.326
	(0.131)***	(0.139)***	(0.170)**	(0.140)***	(0.128)**
$\Delta EDUSH^{H}_{o,t}$	0.916	0.824	1.028	0.888	
	(0.291)***	(0.192)***	(0.226)***	(0.230)***	
$\Delta Gini^{WIDER}_{o,t}$		-0.364		-0.284	
		(0.200)*		(0.182)	
$\Delta GDPPC_{o,t}$		0.955	1.562	-0.685	
		(0.525)*	(0.352)***	(0.782)	
$\Delta Gini^{LIS}_{o,t}$			-0.574		
			(0.204)***		
$\Delta PolityIV_{o,t}$				-0.002	
				(0.003)	
$Conflict_{o,t}$				-0.142	
				(0.057)**	
R^2	0.036	0.066	0.113	0.087	0.243
F-Statistics	71.4	60.9	54.5	60.8	70.7
	B. Dependent variable: Change in Share of Middle Educated				
$RSH_{j,t}$	-0.295	-0.298	-0.315	-0.298	-0.304
	(0.102)***	(0.098)***	(0.124)**	(0.098)***	(0.095)***
$\Delta EDUSH^{M}_{o,t}$	0.520	0.693	0.748	0.668	
	(0.201)***	(0.231)***	(0.246)***	(0.254)***	
$\Delta Gini^{WIDER}_{o,t}$		-0.076		-0.034	
		(0.164)		(0.221)	
$\Delta GDPPC_{o,t}$		1.762	1.661	1.533	
		(0.275)***	(0.264)***	(0.878)*	
$\Delta Gini^{LIS}_{o,t}$			0.115		
			(0.285)		
$\Delta PolityIV_{o,t}$				-0.002	
				(0.003)	
$Conflict_{o,t}$				-0.003	
				(0.061)	
R^2	0.028	0.086	0.093	0.086	0.207
F-Statistics	71.4	60.9	54.4	60.9	70.7
	C. Dependent variable: Change in Share of Low Educated				
$RSH_{j,t}$	-0.046	-0.065	-0.097	-0.065	-0.022
	(0.061)	(0.070)	(0.079)	(0.072)	(0.064)
$\Delta EDUSH^{L}_{o,t}$	0.557	0.560	0.898	0.641	
	(0.356)	(0.291)*	(0.274)***	(0.279)**	
$\Delta Gini^{WIDER}_{o,t}$		0.491		0.386	
		(0.243)**		(0.263)	
$\Delta GDPPC_{o,t}$		-2.720	-3.311	-1.030	
		(0.589)***	(0.376)***	(1.121)	
$\Delta Gini^{LIS}_{o,t}$			0.462		
			(0.276)*		
$\Delta PolityIV_{o,t}$				0.002	
				(0.002)	
$Conflict_{o,t}$				0.148	
				(0.090)*	
R^2	0.020	0.147	0.234	0.166	0.331
F-Statistics	71.4	60.9	54.4	60.9	70.7
Observations	5304	4921	3570	4921	5304
Fixed Effects					$t \times o$

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors in parentheses are robust, clustered by CZ and origin country. Regressions include fixed effects for cantons, origin countries and decades, and are weighted using the total number of recent immigrants from origin country o in destination j at the beginning of the decade. $RSH_{j,t}$ is instrumented with $\widetilde{RSH}_{j,1970}$. $\Delta EDUSH^{e}_{o,t}$ is the decennial change in the share of education group in origin country o for education group $e \in \{H, M, L\}$. $\Delta GDPPC_{o,t}$ and $\Delta Gini^{WIDER}_{o,t}$ represents the decennial change in per capita GDP and Gini index (using WIDER data) in origin country o. $\Delta PolityIV_{o,t}$ represents the decennial change in the policy IV score. $Conflict_{o,t}$ is one if an origin country had at least 25 battle related fatalities in a decade.

Benchmarking the Impact of Educational Supply in Origin Countries and Routine-Biased Demand in Destinations

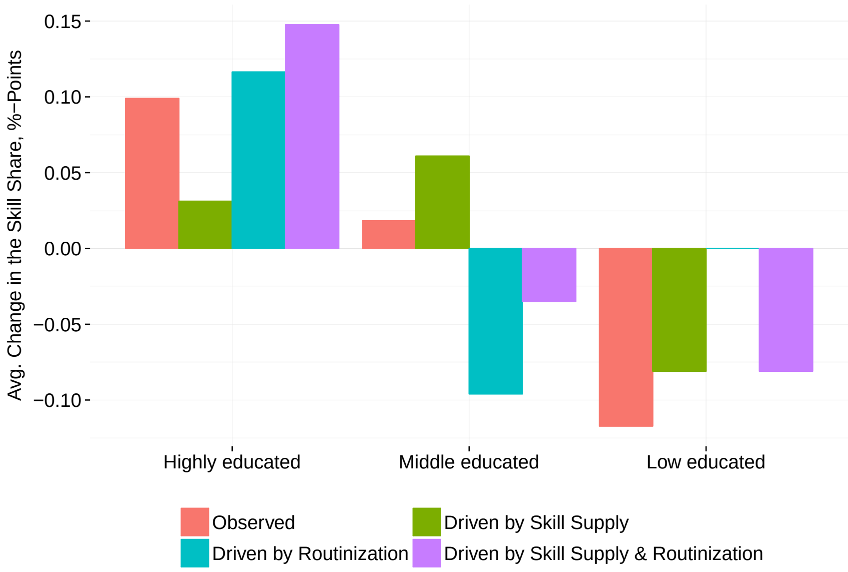
How much of the observed trends in the migrating workers' skill-mix can be accounted for by the two most important factors documented in Table 5.1 (column 4), i.e. skill supply in origins and routine-biased demand shifts in destinations?

As documented above, the share of highly educated new immigrant workers increased from 16% to 46%, or roughly 10 percentage points per decade (ppd). The share of middle educated increased only by 2 ppd and the share of low educated strongly decreased by almost 12 ppd (see blue bars in Figure 5.1). Clearly, part of these changes are just driven by the fact that the education supply changed in the countries where these workers emigrated from. On average, the share of highly, middle and low educated workers in the origin countries increased by 3.5 ppd and 9.1 ppd, respectively, while the share of low educated decreased by 12.6 ppd. Using the estimates of $\Delta EDUSH^{e}_{o,t}$, the red bars show the predicted change in an education group's share using educational supply only. This implies that the share of highly and middle educated new immigrants would increase by 3.1 ppd (0.88×0.035) and 6.1 ppd (0.66×0.091), respectively, whereas the share of low educated would decrease by 8.1 ppd (0.64×-0.126). Hence, the changes in supply clearly underestimate the observed change of highly educated workers and massively overestimate the change in the share of the middle educated group. The prediction is most in line with the decrease of the low educated group.

This highlights the importance of accounting for changes in the relative skill demand, depicted by green bars in Figure 5.1.[39] Adding both the supply and demand effects together (yellow bars), our estimations imply that the share of highly educated recent immigrants increased by 14.7 ppd whereas the share of middle and low educated decreased by 4.5 and 8.1 ppd, respectively.

[39] The coefficients of $RSH_{j,t}$ imply that an average commuting zone with a share of 0.33 in routine employment in 1980 would have experienced an increase in the share of highly educated recent immigrants of 11.6 ppd (0.36×0.33) and a decrease in the share of middle educated immigrants of -9.5 ppd (-0.29×0.33) whereas the impact on low educated workers cannot be distinguished from zero (green bars).

Figure 5.1: Actual and Predicted Changes in the Skill Mix of New Immigrant Workers, 1980 - 2010



Note: The blue bars indicate the actual change in the education group shares 1980, averaged over origin countries and CZs using the number of new immigrants in a cell as weights. Red, green and yellow bars represented the predicted change in the education group share of new immigrants between 1980 and 2010 using the regression model from Table 5.1 (column 4) and the actual, weighted averages of $RSH_{j,1980}$ and $\Delta EDUSH^{e}_{o,1980-2010}$ as explained in the text.

Robustness Tests for Functional Form and Zero Cells

Before analyzing other factors which potentially drive the skill-mix of immigrants, we briefly scrutinize some of the assumptions of our baseline specification.

In Table 5.A.5 in the appendix, we check whether our previous results are sensitive to assumption about the functional form by directly estimating the sorting equation (5.5) of the GH model. In this Table, the dependent variable is either the change in log ratio of highly to middle educated new immigrants (Panel A) or the change in the log ratio of middle to low educated new immigrants (Panel B). Column 1 present the estimates including only the routine share and the supply measure (also as log ratios of the corresponding skill groups). Column 2 presents estimates of Equation (5.5) including the measures for the change in the absolute wage difference between highly and middle educated in the origin countries taken from the LIS data. Column 3 adds all the controls as in Column 4 in Table 5.1. In contrast to the linear utility assumed GH's framework, Columns 4 to 6 present estimates assuming log utility as standard in large parts of the literature (see framework explained in appendix 4). With log utility, immigrant sorting depends on relative wages between skill groups rather than absolute wage differences as in the case of linear utility. In Column 4, the change in the gini coefficient proxies for changes to the returns to skills in the origin countries. In Column 5 and 6 of Panel A, we directly control for changes in the returns to skill of highly educated relative to middle educated (Panel B: middle relative to low educated) using wage measures from the LIS. Column 7, finally, includes origin country-decade fixed effects. The table illustrates several points. First, the routine-biased demand shift positively and significantly affects the relative weight of highly educated over middle educated but not the relative weight of middle over low educated new immigrants irrespective of the assumed functional form. Relative supply also strongly affects the highly over middle educated group but is weaker in the case of the lower educated group. No matter if we assume linear or log utility, the changing relative remuneration of highly educated (compared to middle educated) in the origin countries is negatively associated with the weight of highly educated new immigrants as the models would suggest.[40]

Additionally, Appendix Table 5.A.6 shows that the previous results are robust to dropping all origin country destination cells for which we have replaced missing observations in at least one education group with a zero. We check the sensitivity of our results with respect to violations of the independence of irrelevant alternatives (IIA), by re-estimating our baseline model (Table 5.1 Column 4) 106 times, omitting each destination once. The basic idea (as discussed in GH) is that, if some destinations are perceived as close substitutes by immigrants (violating the IIA assumption), the estimated parameters would be very sensitive to these omissions. Appendix Figure 5.A.1

[40] This complements the findings of GH, who showed for a large cross-section of countries that the sorting equation does not differentiate between the log and the linear utility model in case of destinations with similar productivity levels.

plots the coefficients of the routine share (Panel A) and the change in education supply (Panel B) for each of the 106 different baseline regressions. This shows that the coefficients are very stable across different samples of destinations suggesting that the IIA assumption is not violated in our data.

5.4.2 Robustness to Omitted Pull Factors in Destinations

Routinization may not be the only factor affecting the location decision of immigrants with different educational backgrounds. In Table 5.2, we perform more robustness checks by augmenting our preferred specification from Column 4 in Table 5.1 (reproduced in Column 1) with several other variables that have been considered in the literature on skill-biased technical change and immigrants selection. This table is similarly structured as the previous table with separate regression results for highly (panel A), middle (panel B) and low educated immigrant workers (panel C).

The location choices of new immigrants may be strongly influenced by existing local networks of compatriots (see e.g. Card, 2001). If previous immigrants settled in initially routine-intensive commuting zones and this affected current inflows of their compatriots, the coefficient of our routine measure would be biased. To check the importance of ethnic networks, we include in Column 2 the population share of immigrants from origin country o in destination j in 1970, $IMSH_{j,o,1970}$. Ethnic networks significantly attract low educated workers and reduce the share of middle educated and highly educated new immigrants (not significantly so in the latter case). This is in line with results from Bartel (1989), who finds for the U.S. that more educated immigrants are less likely to settle in cities with a high proportion of a similar ethnic group.

In destinations with initially very low levels of highly educated, subsequent larger upgrades could simply be due to mean reversion, as Michaels, Natraj and Reenen (2014) point out. Column 3 shows that the previous results are not affected to controlling for the initial level of skills of an origin county group of new immigrants in 1970.

Task-offshoring is an important competing explanation for long-run changes to the relative demand for skills (see the literature on wage inequality, e.g. Goos, Manning and Salomons, 2014). In particular, it could be that firms move routine-intensive tasks abroad to lower wage countries while focusing more on higher skill tasks at home. In Column 4, we include a proxy from Blinder and Krueger (2013) on how susceptible local employment is to offshoring.[41] The estimates are very small and statistically insignificant for each education group. This indicates that, as in other countries, technology is the dominant force affecting skill demand, in particular in the case of new immigrants here.[42]

In Column 5, we include the share of employment in manufacturing to check whether the effect of routinization is simply driven by the declining in manufacturing. An initial higher manufacturing share is negatively associated with the growth of high skilled immigrants and positively with middle educated immigrants. The estimated effect of routinization, however, does not change much.

When we include the full set of explanatory variables in the model in Column 6, both the coefficient of the routine share and the education supply remain robustly significant and economically substantial as in previous estimates. In case of the low educated, the coefficient of the routine share becomes slightly more negative and statistically significant. This could point to the fact that low educated immigrants were also employed to a not negligible degree in routine-intensive occupations in 1980 and that routine-biased de-

[41] Blinder and Krueger (2013) provide different measures of the exposure to offshoring, depending on individual worker level characteristics such as occupations or education levels. We use their measure for different education levels assessed by experts. We matched this measure to the education levels in the Swiss Census and computed the average level of susceptibility to offshoring depending on the educational distribution of total employment at the beginning of a decade. The measure was further normalized to have mean zero and standard deviation one across CZs in 1980. See the online appendix for more details.

[42] We also tested other measures of the offshorability of ISCO-occupations taken from Goos, Manning and Salomons (2014) or as defined by Autor and Dorn (2013) for U.S. occupations. These measures were either insignificant or yielded similar results as reported here.

mand might have negatively affected the demand for this group as discussed above.

In the next three columns, we include interactions of canton fixed effects with origin country fixed effects (column 6) or with decade fixed effects (column 7 and 8). Origin country $\times$ canton interactions allow for trends in bilateral migration costs. Including canton-decade fixed effects allows controlling for regional differences (and their changes) in income tax rates or the provision or quality of university education.[43] Even though these are again very demanding specifications, the results for the routine share and the education supply are essentially unchanged. Finally, Column 9 repeats the specification in Column 8 restricting the sample to new immigrants who report non-zero working hours. If estimates for the general population of new immigrants were largely different from those of workers, this would highlight that potentially other local pull-factors, e.g. social security provision, could be important. Again, the estimates coefficients of the routine-biased demand shifts and the supply shift are not statistically different from the previous estimates. This highlights that the location decision of the general new immigrant population is largely similar to those immigrants seeking employment in the same areas.[44]

5.4.3 Robustness with respect to Changes in Immigration Policy

Column 5 of Table 5.1 shows that the effect of routine-biased demand shifts on the skill-mix of immigrants is robust to including origin-country $\times$ decade fixed effects. The latter absorb all immigration policy-relevant variation, as immigration restrictions for newly arriving immigrants residing in Switzer-

[43] Tax rates for residing immigrants vary largely at the level of Cantons, see e.g. Schmidheiny and Slotwinski, 2015. New immigrants with an income level below 120'000 Swiss Francs (roughly \$120'000 at current exchange rates) are subject to source-taxation which is a weighted average of the tax rates of all municipalities within a canton. New immigrant residents above this threshold need to full income taxes on the municipality level like natives. However, as the progressively of tax rates is determined at the Cantonal and Federal level, we are confident that canton $\times$ decade fixed effect absorb all of the relevant variation in local tax rates.

[44] This is not very surprising, as the share of new immigrants reporting positive working hours was 81% in 1980, 79% in 1990, 64% in 2000 and 72 % in 2010.

Table 5.2: Robustness to Omitted Pull Variables, 1980 - 2010, 2SLS

	Population								Workers
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				A. Dependent variable: Change in Share of Highly Educated					
$RSH_{j,t}$	0.362	0.366	0.358	0.388	0.371	0.412	0.335	0.344	0.323
	(0.140)***	(0.139)***	(0.143)**	(0.176)**	(0.129)***	(0.169)**	(0.141)**	(0.136)**	(0.070)***
$\Delta EDUSH^{H}_{o,t}$	0.888	0.887	0.881	0.886	0.889	0.879	0.900	0.867	1.045
	(0.230)***	(0.230)***	(0.229)***	(0.226)***	(0.230)***	(0.226)***	(0.234)***	(0.183)***	(0.282)***
$IMSH_{j,o,1970}$		-0.112				-0.116			
		(0.086)				(0.080)			
$EDUSH^{H}_{j,o,1970}$			-0.017			-0.026			
			(0.016)			(0.017)			
$EDUSH^{M}_{j,o,1970}$			0.020			0.018			
			(0.015)			(0.014)			
$OFFSH_{j,t}$				-0.005		-0.007			
				(0.008)		(0.008)			
$(Manu/Emp)_{j,t}$					-0.092	-0.115			
					(0.023)***	(0.023)***			
R^2	0.087	0.087	0.096	0.087	0.088	0.097	0.192	0.131	0.124
F-Statistics	60.8	60.9	55.9	61.4	103.7	96.2	51.8	58.6	66.7
				B. Dependent variable: Change in Share of Middle Educated					
$RSH_{j,t}$	-0.298	-0.293	-0.276	-0.278	-0.303	-0.265	-0.289	-0.287	-0.301
	(0.098)***	(0.096)***	(0.099)***	(0.113)**	(0.096)***	(0.112)**	(0.096)***	(0.097)***	(0.060)***
$\Delta EDUSH^{M}_{o,t}$	0.668	0.667	0.680	0.675	0.670	0.685	0.668	0.549	0.818
	(0.254)***	(0.254)***	(0.252)***	(0.251)***	(0.256)***	(0.248)***	(0.261)**	(0.190)***	(0.308)***
$IMSH_{j,o,1970}$		-0.143				-0.124			
		(0.082)*				(0.096)			
$EDUSH^{H}_{j,o,1970}$			0.002			0.006			
			(0.019)			(0.018)			
$EDUSH^{M}_{j,o,1970}$			-0.018			-0.013			
			(0.014)			(0.013)			
$OFFSH_{j,t}$				-0.004		-0.002			
				(0.006)		(0.007)			
$(Manu/Emp)_{j,t}$					0.054	0.052			
					(0.031)*	(0.033)			
R^2	0.086	0.087	0.098	0.087	0.086	0.100	0.181	0.148	0.106
F-Statistics	60.9	60.9	56.0	61.5	103.8	96.3	51.8	58.6	66.7
				C. Dependent variable: Change in Share of Low Educated					
$RSH_{j,t}$	-0.065	-0.075	-0.083	-0.111	-0.069	-0.148	-0.047	-0.059	-0.023
	(0.072)	(0.070)	(0.078)	(0.081)	(0.065)	(0.073)**	(0.077)	(0.072)	(0.068)
$\Delta EDUSH^{L}_{o,t}$	0.641	0.638	0.653	0.654	0.640	0.663	0.660	0.622	0.721
	(0.279)**	(0.280)**	(0.272)**	(0.282)**	(0.279)**	(0.275)**	(0.288)**	(0.307)**	(0.316)**
$IMSH_{j,o,1970}$		0.260				0.244			
		(0.087)***				(0.093)***			
$EDUSH^{H}_{j,o,1970}$			0.015			0.020			
			(0.023)			(0.023)			
$EDUSH^{M}_{j,o,1970}$			-0.003			-0.006			
			(0.010)			(0.010)			
$OFFSH_{j,t}$				0.008		0.009			
				(0.006)		(0.006)			
$(Manu/Emp)_{j,t}$					0.038	0.062			
					(0.036)	(0.028)**			
R^2	0.166	0.167	0.187	0.168	0.166	0.190	0.255	0.199	0.120
F-Statistics	60.9	60.9	55.9	61.4	103.8	96.2	51.8	58.6	66.7
Observations	4921	4921	4166	4921	4921	4166	4921	4921	4213
Fixed Effects							$c \times o$	$c \times t$	$c \times t$

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors in parentheses are robust, clustered by CZ and origin country. Regressions include fixed effects for cantons, origin countries and decades, and are weighted using the total number of recent immigrants from origin country o in destination j at the beginning of the decade. $RSH_{j,t}$ is instrumented with $\widetilde{RSH}_{j,1970}$. $\Delta EDUSH^{e}_{o,t}$ is the decennial change in the share of education group in origin country o for education group $e \in \{H, M, L\}$. $EDUSH^{H,M}_{j,o,1970}$ is the share of highly and middle educated new immigrants in 1970, respectively. $IMSH_{j,o,1970}$ is the population share of immigrants from origin country o in destination j in 1970. $OFFSH_{j,t}$ is a measure of the potential to offshoring a CZ's employment at the beginning of a decade as defined in Blinder and Krueger (2013). $(Manu/Emp)_{j,t}$ is the employment share of manufacturing in a CZ at the beginning of a decade. All regressions include controls for the changes of GDP per capita, the GINI index (WIDER), the polity IV measure and a dummy for whether a country was involved in a conflict.

land varied by origin country and decade (as we describe next). One remaining question, however, is whether we can also corroborate the robustness of the effect of skill supply from origin countries by including controls for immigration policies. Although the variation across three decades and essentially two origin country groups does not allow identifying the causal effects of immigration policies with full rigor, we use the estimates of the immigration policy controls to get a sense of the relative importance of immigration policies vis-à-vis the other forces.

Immigration Policies 1980-2010

Immigration policy in Switzerland differed for individuals from different origin countries and changed a couple of times between 1980 and 2010 (Appendix 5 provides a more detailed description). During the 1980s, all immigrants had to apply for a residency permit. Permits were subject to global, yearly quotas set by the federal government. Immigration offices would grant permits only if no equally qualified native could be found for a given job vacancy. A first change to this regime came at the beginning of the 1990s when immigration offices began to favor migrants from EU17 and EFTA countries while migrants from other countries were required to be "highly qualified".[45] A second policy change came through the enactment of the Free Movement of Persons (FMP) policy with the EU after 2002. The FMP policy gradually abolished all immigration restrictions. First for EU17 countries and a few years later for eastern European member states (EU8 and EU2) while quotas and qualification requirements for non-European immigrants remained in place.[46]

[45] "High qualification" mainly meant "managers, specialists and qualified professionals" but was deliberately defined broadly to include also investors, workers who specialized in the construction of certain facilities or tunnels or workers with "special knowledge" (Bundesrat, 1991, 2002). EU17 countries include Belgium, Germany, France, Italy, Luxembourg, Netherlands, Denmark, Ireland, United Kingdom, Greece, Portugal, Spain, Finland, Austria, Sweden, Cyprus and Malta, while Iceland, Norway and Liechtenstein are EFTA countries.

[46] EU8 includes Estonia, Latvia, Lithuania, Poland, Slovakia, Slovenia, Hungary and the Czech Republic. EU2 includes Romania and Bulgaria.

Accounting for the Impact of Immigration Policies

In the framework of GH, changes to the immigration policy alter, ceteris paribus, the immigration costs in the sorting equation (5.4), $\left(g^H - g^M\right)_{j,o,t}$. As we lack direct information on education specific immigration costs, we can only approximate their effect *indirectly* by analyzing whether the skill-mix of immigrants *responds differentially* depending on country- and decade specific entry restrictions.

In contrast to the policy in the 1990s, the FMP policy was not particularly sought for by the Swiss government but requested by the EU to grant access to a larger package of bilateral agreements including reduction of trade barriers (Bundesrat, 1999). Thus, the implementation of this policy is less likely to be triggered by pre-existing demand or supply-trends.[47] We measure its effect by augmenting our baseline specification (5.6) in the following way:

$$\begin{aligned}\Delta EDUSH^e_{j,o,t} \;=\;& \beta^e_1 RSH_{j,t} + \beta^e_2 \Delta X^e_{o,t} + \beta^e_3 \Delta EDUSH^e_{o,t} \\ & + \beta^e_4 1(o = EU) \times 1(t = 2000) + \alpha_c + \alpha_o + \alpha_t + \epsilon_{j,o,t}\end{aligned} \tag{5.10}$$

where $1(o = EU) \times 1(t = 2000)$ is one for countries affected by the free movement policy in the decade 2000-2010 and zero otherwise. In this specification, β^E_4 measures the degree to which education shares of immigrants from EU countries changed differentially between 2000 and 2010 relative to other countries and previous decades. We include the full set of controls used Column 6 of Table 5.2 without the education group shares in 1970, which are not available for all origin-destination pairs.

Table 5.3 shows the results of estimating versions of specification (5.10) for each education group (panel A, B and C). Column 1 repeats the specification in Table 5.2 Column 6. The specification in Column 2 includes a dummy for all EU/EFTA countries interacted with a dummy for the decade 2000-2010.

[47] In contrast, with the introduction of the policy in the early 1990s, the government responded, at least partly, to requests for business to facilitate the hiring of qualified workers while respecting concerns of the population about cultural differences to immigrant from earlier episodes.

The results indicate that the increase in share of highly educated immigrants was roughly 7 percentage points lower among EU countries compared to other countries and previous decades. On the other hand, the increase in the share of low educated was higher for immigrants with EU-origin while the effect for middle educated immigrants is positive but not significant. Importantly, the estimates for both the routine-biased demand effect and the supply effect are both not statistically different from their estimate in Column 1. In Column 3, we estimate two separate effects for immigrants from EU17/EFTA and from EU10 countries (Hungary, Poland and migrants from the former Czechoslovakia plus Romania in our sample). These estimates indicate that the overall negative effect for EU countries is a combination of two heterogenous effects: For immigrants from EU17 countries, the lower increase of highly educated was compensated mostly by low (significant) and middle (not significant) educated individuals. In the case of the EU10 countries, there seemed to be a rebalancing from the middle to the low educated group with a slight (but not significant) gain at the top.

In Column 4, we include an interaction for each of the three origin-country groups with differential policy treatments (EU17, EU10 and Non-EU) with each of the two decades, the 1990s and 2000s. The coefficient of each interaction effect indicates how the education-mix of a group changed relative to its own change during the 1980s.[48] As the Swiss government changed immigration restrictions across decades only across these three groups, these interactions should absorb all the policy-relevant variation in the data. Again, both the estimates for the routine-biased demand shift and the supply effect are not statistically different from their counterparts in Column 1. In addition, we can now gauge the effect of the FMP policy by comparing the difference in change of the skill-mix of immigrants from EU17/EFTA and Non-EU countries only across the two decades from 1990 to 2010. During this time, only the EU17/EFTA group received a change in the immigration

[48] This shows, for instance, that the share of highly educated immigrants from Non-EU countries increased more in the 1990s and 2000s at the expense of middle and low educated. Also the share of highly educated increased more for EU17/EFTA immigrants at the expense of middle educated during the 1990s relative to the 1980s but not in the 2000s.

policy (from quotas to free movement) while the immigration restrictions for the control group of Non-EU countries remained unchanged (qualification restrictions and quotas). The last two rows in each panel report this difference-in-difference effect and its respective p-value. This shows that the share of highly educated immigrants increased roughly 7 percentage points less and significantly so, among EU17/EFTA immigrants. Meanwhile, middle and low educated immigrants have increased their share in equal terms (roughly 3.5 percentage points each) but not significantly so.[49]

In sum, including all the interactions absorbing potential policy effects did not significantly alter our previous estimates for the routine-biased policy effect and the effect of education supply in the origin countries. This is reassuring and indicates that, overall, these two forces together are key to understand the change in the skill-mix of immigrants from a long-run perspective.

In addition, the results also indicate that policy effects have the potential to moderate the long-run trends induced by market forces. For instance, the abolition of immigration restrictions for EU workers, seems to have reduced the migration costs more for low and middle educated individuals increasing their shares relative to the Non-EU group. Notably, this results points in similar direction and is of similar magnitude as findings of Beerli and Peri (2015), see Table 5 of their paper and reproduced in the Appendix 2. Beerli and Peri (2015) use a different data set of new immigrants (including cross-border workers) employed in the private sector which does not allow to distinguish between origin countries but is available biannually from 1994 to 2010. Exploiting the finer time resolution and the fact that the free move-

[49] Formally, this effect is calculated as the difference of the estimated coefficients between EU/EFTA and Non-EU countries in the 2000s and the 1990s, i.e. $(\beta\,[1(o = EU17, EFTA) \cdot 1(t = 2000)] - \beta\,[1(o = NONEU) \cdot 1(t = 2000)]) - (\beta\,[1(o = EU17, EFTA) \cdot 1(t = 1990)] - \beta\,[1(o = NONEU) \cdot 1(t = 1990)])$. In case of immigrants from EU10 countries, the comparison with immigrants from Non-EU countries indicates that the policy still seems to have reallocated immigrants away from the middle educated group (-11 percentage points, not significant) but now the gains at the top increased (+7 percentage points, not significant) while the gain at the bottom is smaller and turned insignificant (+4 percentage points). As the variation for this group comes only from four countries, these estimates have to be interpreted with caution and are omitted from the table.

Table 5.3: The Effect of the Free Movement Policy on the Skill-Mix of Immigrants, 2SLS

	(1)	(2)	(3)	(4)
A. Dependent variable: Change in Share of Highly Educated				
$RSH_{j,t}$	0.413	0.401	0.400	0.399
	(0.166)**	(0.161)**	(0.161)**	(0.161)**
$\Delta EDUSH^{H}_{o,t}$	0.885	0.814	0.782	0.955
	(0.227)***	(0.180)***	(0.166)***	(0.238)***
$1(o = EU, EFTA) \times 1(t = 2000)$		-0.073		
		(0.029)**		
$1(o = EU17, EFTA) \times 1(t = 2000)$			-0.088	0.045
			(0.030)***	(0.031)
$1(o = EU17, EFTA) \times 1(t = 1990)$				0.124
				(0.036)***
$1(o = EU10) \times 1(t = 2000)$			0.035	0.117
			(0.046)	(0.104)
$1(o = EU10) \times 1(t = 1990)$				0.054
				(0.099)
$1(o = NONEU) \times 1(t = 2000)$				0.182
				(0.063)***
$1(o = NONEU) \times 1(t = 1990)$				0.195
				(0.059)***
R^2	0.088	0.097	0.101	0.229
DID Effect of AFMP (EU17-Non-EU)				-0.066
P-Value of DID Effect of AFMP				0.049
B. Dependent variable: Change in Share of Middle Educated				
$RSH_{j,t}$	-0.284	-0.279	-0.278	-0.277
	(0.112)**	(0.107)***	(0.108)**	(0.108)**
$\Delta EDUSH^{M}_{o,t}$	0.673	0.679	0.591	0.714
	(0.250)***	(0.240)***	(0.229)**	(0.240)***
$1(o = EU, EFTA) \times 1(t = 2000)$		0.026		
		(0.030)		
$1(o = EU17, EFTA) \times 1(t = 2000)$			0.046	0.023
			(0.029)	(0.043)
$1(o = EU17, EFTA) \times 1(t = 1990)$				-0.087
				(0.040)**
$1(o = EU10) \times 1(t = 2000)$			-0.133	-0.241
			(0.070)*	(0.079)***
$1(o = EU10) \times 1(t = 1990)$				-0.203
				(0.087)**
$1(o = NONEU) \times 1(t = 2000)$				-0.068
				(0.053)
$1(o = NONEU) \times 1(t = 1990)$				-0.142
				(0.047)***
R^2	0.088	0.089	0.098	0.243
DID Effect of AFMP (EU17-Non-EU)				0.036
P-Value of DID Effect of AFMP				0.182
C. Dependent variable: Change in Share of Low Educated				
$RSH_{j,t}$	-0.131	-0.122	-0.123	-0.123
	(0.067)*	(0.066)*	(0.065)*	(0.065)*
$\Delta EDUSH^{L}_{o,t}$	0.651	0.610	0.592	0.685
	(0.283)**	(0.290)**	(0.289)**	(0.295)**
$1(o = EU, EFTA) \times 1(t = 2000)$		0.052		
		(0.021)**		
$1(o = EU17, EFTA) \times 1(t = 2000)$			0.047	-0.069
			(0.020)**	(0.040)*
$1(o = EU17, EFTA) \times 1(t = 1990)$				-0.046
				(0.029)
$1(o = EU10) \times 1(t = 2000)$			0.096	0.105
			(0.062)	(0.129)
$1(o = EU10) \times 1(t = 1990)$				0.123
				(0.139)
$1(o = NONEU) \times 1(t = 2000)$				-0.111
				(0.051)**
$1(o = NONEU) \times 1(t = 1990)$				-0.048
				(0.070)
R^2	0.169	0.173	0.173	0.188
DID Effect of AFMP (EU17-Non-EU)				0.040
P-Value of DID Effect of AFMP				0.169

Note: The table reports the results for augmented versions of specification (4) in Table 5.1 (see the notes to Table 5.1). $1(o = \text{country group}) \times 1(t = \text{decade})$ represents an interaction term of a country group (EU+EFTA, EU17+EFTA, EU10, Non-EU) with the dummy for the indicated decade. "DID Effect of FMP (EU17-Non-EU)" represents the difference-in-difference effect of the FMP policy on EU17 immigrants, calculated as $(\beta\,[1(o = EU17, EFTA) \cdot 1(t = 2000)] - \beta\,[1(o = NONEU) \cdot 1(t = 2000)]) - (\beta\,[1(o = EU17, EFTA) \cdot 1(t = 1990)] - \beta\,[1(o = NONEU) \cdot 1(t = 1990)])$, separately for each skill group.

ment policy was first implemented first only in one region in Switzerland for cross-border workers and later in the rest of the country, they show that the free movement policy increased the share of low educated and reduced the share middle educated among new immigrant workers.

Furthermore, our results complement the findings of Huber and Bock-Schappelwein (2014), who show that Austria's accession to the European Economic Area (EEA) in 1994 reduced the share of low educated permanent immigrants from the EEA compared to other countries. In contrast to the situation in Switzerland, Huber and Bock-Schappelwein (2014) point out that Austria had very low returns to education and immigrants were mostly negatively selected prior to the EEA accession. Thus, their results can be reconciled with ours from the perspective of Borjas (1987)'s immigration selection model in which lowering immigration restrictions increases the incentives for the 'marginal immigrant' to enter the country. In the Swiss case of positive selection prior to the opening, the marginal immigrant is in the lower part of the skill distribution whereas she is in the higher part of the skill-distribution in the case of negative selection in Austria.

5.5 Conclusion

Between 1980 and 2010, the share of immigrants with tertiary education rose on average 19 percentage points in OECD countries. Astonishingly, this increase among migrants was about twice as large as the corresponding gains in their origin country populations. In these countries, educational upgrading resulted primarily in a higher share of secondary educated while the population share with primary education or less decreased strongly. Yet, while the latter also decreased strongly among immigrants, there were only modest gains in the middle. This means that immigrants have become more positively selected throughout this period. As highly educated immigrants are often seen as net-contributors to receiving economies and many countries adopted policies to attract or facilitate immigration of highly skilled workers, it is important to understand the drivers of these trends.

In this paper, we analyze the factors driving the skill-mix of new immigrant workers using a framework suggested by Grogger and Hanson (2011). According to this framework, the selection of immigrants changes if the demand for workers with different skills changes differentially in the destinations than in the origin countries, or if skill-related migration costs change, e.g. due to selective immigration policies. We apply this framework to Switzerland as a country with high inflows of foreign workers, which also experienced changes in their skill-mix paralleling those in other OECD countries. We use variation in the skill-mix of immigrants from 30 origin countries which newly arrive and settle across different Swiss local labor markets. This allows us exploiting that these local labor markets were exposed to different skill-biased demand shifts due to their industrial composition in 1970. Regions with a higher initial share of workers employed in routine occupations experienced stronger adoption of computer capital. This increased (decreased) the demand for highly (middle) educated workers raising both wages and employment of highly relative to middle educated workers. Consequentially, these routine biased demand shifts should also lead to an increase of highly educated over middle educated immigrants. Our findings suggest that changes of education supply in origin countries and shifts to the relative skill demand stand out as the two most important drivers. Yet, supply alone predicts only a modest increase in the case of highly educated workers and a large increase of middle educated workers. In contrast, routine-biased demand shifts are crucial to explain the sharp increase in highly educated workers and the mere stabilization of the share of middle educated immigrants. These effects are very robust to a wide range of checks. Most importantly, they are not affected by different ways of controlling for changes to earnings and inequality or other characteristics of the origin countries of immigrants. Also other pull factors, such as other types of (skill-biased) demand shifts, ethnic networks or institutional differences between destinations, do not affect the estimated effects. Finally, we show that immigration policies have contributed only modestly to the observed changes in the skill-mix of new immigrants. Again, the estimates of the supply and demand effects are not changed if we control for policy effects.

Appendix

5.A Additional Figures and Tables

Table 5.A.1: Employment Shares of Occupation Groups by Nationality, 1980 - 2010

		Occupation Group Shares				Change (%-Points)
ISCO-88 Code	Occupation	1980	1990	2000	2010	1980-2010
		A. Total Workforce				
2	Professionals	0.088	0.108	0.147	0.171	0.082
1	Managers	0.057	0.098	0.107	0.116	0.059
3	Technicians	0.129	0.193	0.206	0.219	0.090
4	Clerks	0.193	0.150	0.138	0.104	-0.089
7	Craft	0.273	0.215	0.169	0.149	-0.124
8	Operators/Assemblers	0.081	0.062	0.052	0.044	-0.037
9	Elementary Occ.	0.045	0.042	0.040	0.052	0.007
5	Service/Sales	0.134	0.131	0.142	0.146	0.012
		B. Natives				
2	Professionals	0.088	0.113	0.151	0.174	0.086
1	Managers	0.062	0.106	0.111	0.114	0.051
3	Technicians	0.140	0.210	0.219	0.239	0.099
4	Clerks	0.213	0.167	0.147	0.114	-0.099
7	Craft	0.247	0.196	0.162	0.142	-0.105
8	Operators/Assemblers	0.074	0.055	0.046	0.038	-0.036
9	Elementary Occ.	0.042	0.033	0.032	0.038	-0.004
5	Service/Sales	0.134	0.120	0.132	0.141	0.007
		C. Recent Immigrants				
2	Professionals	0.090	0.089	0.228	0.202	0.112
1	Managers	0.027	0.046	0.129	0.145	0.118
3	Technicians	0.078	0.130	0.175	0.164	0.087
4	Clerks	0.074	0.057	0.069	0.064	-0.010
7	Craft	0.417	0.290	0.112	0.150	-0.267
8	Operators/Assemblers	0.075	0.061	0.032	0.040	-0.036
9	Elementary Occ.	0.050	0.078	0.060	0.088	0.037
5	Service/Sales	0.188	0.249	0.195	0.148	-0.040

Note: Employment shares (in full time equivalents) of ISCO main occupation groups (omitting agriculture). Occupations are ranked in a descending order by their median wage taken from the Swiss Labour Force Survey 1991-1993 (pooled). Employment data from Swiss Census, 1980-2010

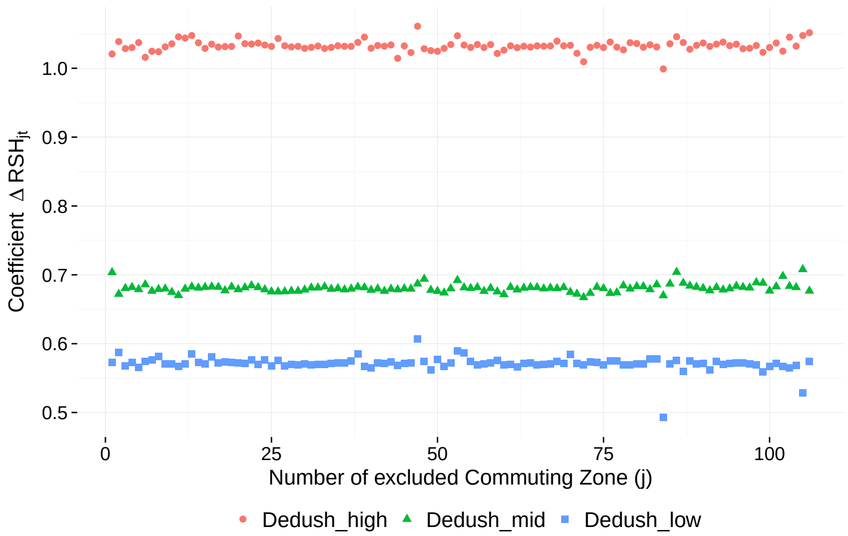
Table 5.A.2: Average Skill Group Shares: New Immigrants in Switzerland by Origin Country, Origin Country Population

			Education groups, new imm. workers, in destinations						Education groups, origin country population					
	Share new imm.		Share 1980			Change share 1980-2010			Share 1980			Change share 1980-2010		
Origin Country	1980	2010	high	middle	low	high	middle	low	high	middle	low	high	middle	low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
EU countries														
Austria	0.027	0.023	0.191	0.459	0.350	0.345	-0.099	-0.246	0.027	0.480	0.493	0.114	0.120	-0.235
Belgium	0.009	0.009	0.502	0.296	0.202	0.223	-0.136	-0.087	0.113	0.426	0.461	0.130	0.126	-0.256
Germany	0.090	0.315	0.320	0.450	0.230	0.213	-0.055	-0.157	0.057	0.201	0.742	0.121	0.517	-0.638
Denmark	0.003	0.004	0.315	0.394	0.291	0.437	-0.206	-0.232	0.135	0.370	0.494	0.054	0.030	-0.084
Spain	0.125	0.026	0.037	0.110	0.854	0.460	0.073	-0.533	0.081	0.258	0.661	0.162	0.251	-0.413
Finland	0.005	0.003	0.445	0.447	0.108	0.322	-0.265	-0.056	0.108	0.351	0.541	0.132	0.082	-0.214
France	0.070	0.095	0.273	0.410	0.317	0.386	-0.219	-0.168	0.077	0.238	0.685	0.125	0.400	-0.525
Greece	0.007	0.005	0.193	0.248	0.559	0.572	-0.110	-0.462	0.086	0.254	0.660	0.153	0.220	-0.373
Hungary	0.007	0.011	0.346	0.422	0.232	0.156	-0.023	-0.133	0.068	0.306	0.626	0.075	0.495	-0.569
Italy	0.222	0.077	0.068	0.168	0.763	0.372	0.127	-0.498	0.033	0.397	0.570	0.060	0.271	-0.331
Netherlands	0.015	0.013	0.341	0.445	0.214	0.414	-0.302	-0.112	0.107	0.599	0.295	0.123	0.047	-0.170
Poland	0.005	0.018	0.539	0.333	0.128	-0.049	0.038	0.011	0.046	0.395	0.559	0.101	0.237	-0.339
Portugal	0.073	0.133	0.019	0.104	0.877	0.062	0.006	-0.067	0.035	0.159	0.806	0.075	0.189	-0.264
Romania	0.005	0.008	0.604	0.273	0.124	-0.074	0.021	0.053	0.046	0.555	0.399	0.072	0.196	-0.268
Sweden	0.007	0.006	0.488	0.355	0.157	0.244	-0.196	-0.048	0.134	0.487	0.379	0.085	0.185	-0.270
Czech Rep.+Slovakia	0.009	0.017	0.388	0.450	0.162	-0.073	0.138	-0.065	0.051	0.472	0.477	0.062	0.255	-0.317
United Kingdom	0.033	0.043	0.465	0.333	0.202	0.336	-0.216	-0.121	0.090	0.416	0.493	0.136	0.053	-0.189
Non-EU countries														
Canada	0.009	0.008	0.464	0.351	0.185	0.295	-0.207	-0.088	0.257	0.536	0.207	0.095	0.053	-0.148
Chile	0.005	0.002	0.258	0.272	0.469	0.205	-0.089	-0.116	0.071	0.337	0.592	0.205	0.129	-0.334
China	0.002	0.010	0.492	0.091	0.416	0.042	0.065	-0.107	0.009	0.336	0.655	0.071	0.273	-0.345
Algeria	0.004	0.003	0.373	0.269	0.359	-0.001	-0.022	0.023	0.013	0.139	0.848	0.105	0.266	-0.371
Israel	0.004	0.002	0.359	0.393	0.248	0.296	-0.184	-0.112	0.198	0.464	0.338	0.110	-0.022	-0.087
India	0.005	0.010	0.451	0.208	0.341	0.407	-0.144	-0.263	0.023	0.187	0.790	0.052	0.203	-0.255
Iran	0.005	0.004	0.441	0.282	0.277	0.073	-0.016	-0.056	0.021	0.201	0.778	0.161	0.292	-0.453
Japan	0.005	0.005	0.724	0.171	0.106	0.070	-0.029	-0.041	0.156	0.455	0.390	0.219	0.001	-0.220
Balkan countries	0.156	0.093	0.023	0.131	0.846	0.090	0.208	-0.298	0.054	0.322	0.624	0.055	0.272	-0.328
Tunisia	0.003	0.006	0.178	0.384	0.438	0.171	-0.156	-0.014	0.018	0.152	0.829	0.099	0.223	-0.323
Turkey	0.049	0.021	0.048	0.113	0.839	0.140	0.075	-0.216	0.030	0.141	0.828	0.084	0.227	-0.312
USA	0.028	0.027	0.699	0.167	0.134	0.133	-0.072	-0.061	0.299	0.642	0.059	0.218	-0.188	-0.031
Vietnam	0.016	0.002	0.049	0.145	0.806	0.264	0.172	-0.436	0.009	0.409	0.582	0.060	-0.118	0.059
Mean			0.162	0.222	0.616	0.297	0.052	-0.349	0.064	0.316	0.620	0.105	0.274	-0.379

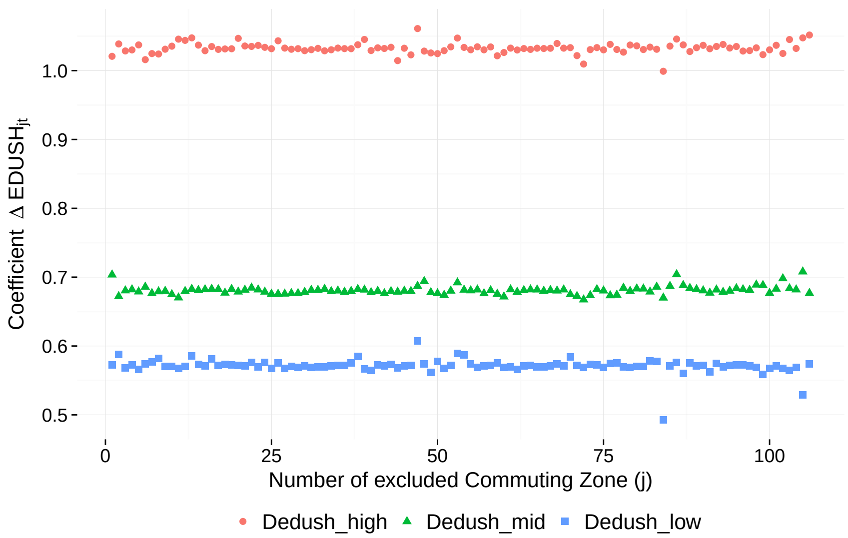
Note: Swiss Census 1980 and 2010 and Barro and Lee (2013).

Figure 5.A.1: Plot of Coefficients, Omitting Each Destination Once

A. Coefficient of $RSH_{j,t}$



B. Coefficient of $\Delta EDUSH^{e}_{o,t}$



Note: The graphs show the estimated coefficient of $RSH_{j,t}$ and $\Delta EDUSH_{o,t}$ of 106 separate regressions for each education group, omitting each commuting zone once.

Table 5.A.3: Determinants of the Change in Skill Group Shares, 1980 - 2010, OLS

	(1)	(2)	(3)	(4)	(5)
A. Dependent variable: Change in Share of Highly Educated					
$RSH_{j,t}$	0.154	0.202	0.211	0.195	0.157
	(0.056)***	(0.057)***	(0.064)***	(0.056)***	(0.055)***
$\Delta EDUSH^{H}_{o,t}$	0.914	0.822	1.037	0.889	
	(0.295)***	(0.194)***	(0.233)***	(0.231)***	
$\Delta Gini^{WIDER}_{o,t}$		-0.365		-0.282	
		(0.205)*		(0.186)	
$\Delta GDPPC_{o,t}$		0.954	1.562	-0.702	
		(0.528)*	(0.353)***	(0.786)	
$\Delta Gini^{LIS}_{o,t}$			-0.587		
			(0.211)***		
$\Delta PolityIV_{o,t}$				-0.002	
				(0.003)	
$Conflict_{o,t}$				-0.142	
				(0.058)**	
R^2	0.041	0.069	0.118	0.090	0.246
B. Dependent variable: Change in Share of Middle Educated					
$RSH_{j,t}$	-0.199	-0.184	-0.209	-0.186	-0.193
	(0.064)***	(0.063)***	(0.074)***	(0.063)***	(0.061)***
$\Delta EDUSH^{M}_{o,t}$	0.522	0.696	0.752	0.672	
	(0.202)***	(0.232)***	(0.246)***	(0.256)***	
$\Delta Gini^{WIDER}_{o,t}$		-0.075		-0.035	
		(0.167)		(0.224)	
$\Delta GDPPC_{o,t}$		1.764	1.661	1.546	
		(0.273)***	(0.264)***	(0.876)*	
$\Delta Gini^{LIS}_{o,t}$			0.120		
			(0.285)		
$\Delta PolityIV_{o,t}$				-0.001	
				(0.003)	
$Conflict_{o,t}$				-0.003	
				(0.061)	
R^2	0.030	0.087	0.094	0.088	0.208
C. Dependent variable: Change in Share of Low Educated					
$RSH_{j,t}$	0.046	-0.015	-0.007	-0.009	0.036
	(0.064)	(0.069)	(0.091)	(0.071)	(0.064)
$\Delta EDUSH^{L}_{o,t}$	0.554	0.559	0.897	0.639	
	(0.357)	(0.293)*	(0.277)***	(0.282)**	
$\Delta Gini^{WIDER}_{o,t}$		0.492		0.386	
		(0.243)**		(0.264)	
$\Delta GDPPC_{o,t}$		-2.720	-3.312	-1.025	
		(0.590)***	(0.376)***	(1.125)	
$\Delta Gini^{LIS}_{o,t}$			0.467		
			(0.279)*		
$\Delta PolityIV_{o,t}$				0.002	
				(0.002)	
$Conflict_{o,t}$				0.148	
				(0.090)*	
R^2	0.021	0.147	0.235	0.166	0.332
Observations	5304	4921	3570	4921	5304
Fixed Effects					$t \times o$

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors in parentheses are robust, clustered by CZ and origin country. Regressions include fixed effects for cantons, origin countries and decades, and are weighted using the total number of recent immigrants from origin country o in destination j at the beginning of the decade. $\Delta EDUSH^{e}_{o,t}$ is the decennial change in the share of education group in origin country o for education group $e \in \{H, M, L\}$. $\Delta GDPPC_{o,t}$ and $\Delta Gini^{WIDER,LIS}_{o,t}$ represent the decennial change in per capita GDP and Gini index (either using WIDER or LIS data) in origin country o. $\Delta PolityIV_{o,t}$ represents the decennial change in the policy IV score. $Conflict_{o,t}$ is one if an origin country had at least 25 battle related fatalities in a decade.

Table 5.A.4: Determinants of the Change in Skill Group Shares, 1980 - 2010, 2SLS, Controls in Levels

	(1)	(2)	(3)	(4)	(5)
	A. Dependent variable: Change in Share of Highly Educated				
$RSH_{j,t}$	0.340	0.370	0.414	0.359	0.326
	(0.131)***	(0.141)***	(0.170)**	(0.140)**	(0.128)**
$\Delta EDUSH^{H}_{o,t}$	0.916	0.896	1.163	0.901	
	(0.291)***	(0.252)***	(0.275)***	(0.235)***	
$Gini^{WIDER}_{o,t}$		0.286		-0.051	
		(0.209)		(0.147)	
$\ln GDPPC_{o,t}$		-0.142	-0.371	0.102	
		(0.127)	(0.136)***	(0.078)	
$Gini^{LIS}_{o,t}$			0.430		
			(0.158)***		
$PolityIV_{o,t}$				0.004	
				(0.002)*	
$Conflict_{o,t}$				-0.098	
				(0.035)***	
R^2	0.036	0.051	0.097	0.082	0.243
F-Statistics	71.4	60.9	54.4	60.8	70.7
	B. Dependent variable: Change in Share of Middle Educated				
$RSH_{j,t}$	-0.295	-0.285	-0.313	-0.293	-0.304
	(0.102)***	(0.102)***	(0.126)**	(0.098)***	(0.095)***
$\Delta EDUSH^{M}_{o,t}$	0.520	0.600	0.746	0.598	
	(0.201)***	(0.217)***	(0.219)***	(0.269)**	
$Gini^{WIDER}_{o,t}$		0.596		0.335	
		(0.192)***		(0.233)	
$\ln GDPPC_{o,t}$		-0.285	-0.343	-0.115	
		(0.120)**	(0.171)**	(0.103)	
$Gini^{LIS}_{o,t}$			0.035		
			(0.265)		
$PolityIV_{o,t}$				0.004	
				(0.004)	
$Conflict_{o,t}$				-0.049	
				(0.064)	
R^2	0.028	0.066	0.063	0.080	0.207
F-Statistics	71.4	60.9	54.4	60.9	70.7
	C. Dependent variable: Change in Share of Low Educated				
$RSH_{j,t}$	-0.046	-0.086	-0.103	-0.069	-0.022
	(0.061)	(0.072)	(0.077)	(0.073)	(0.064)
$\Delta EDUSH^{L}_{o,t}$	0.557	0.556	1.082	0.644	
	(0.356)	(0.369)	(0.362)***	(0.320)**	
$Gini^{WIDER}_{o,t}$		-0.959		-0.368	
		(0.354)***		(0.290)	
$\ln GDPPC_{o,t}$		0.422	0.753	0.023	
		(0.236)*	(0.270)***	(0.103)	
$Gini^{LIS}_{o,t}$			-0.466		
			(0.206)**		
$PolityIV_{o,t}$				-0.006	
				(0.003)**	
$Conflict_{o,t}$				0.166	
				(0.067)**	
R^2	0.020	0.090	0.165	0.162	0.331
F-Statistics	71.4	60.9	54.4	60.8	70.7
Observations	5304	4973	3570	4973	5304
Fixed Effects					$t \times o$

Note: See the notes to Table 5.1. This Table reports the results for regressions with the per capita GDP, the Gini index (either using WIDER or LIS data) and the policy IV score taken in levels.

Table 5.A.5: Sorting Equation (5), 1980 - 2010, 2SLS

		Linear utility model		Log utility model			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Dependent variable: Change in Ln(Highly Educated/Middle Educated)							
$RSH_{j,t}$	1.482	1.777	1.725	1.529	1.749	1.720	1.462
	(0.606)**	(0.754)**	(1.064)	(0.667)**	(0.846)**	(0.804)**	(0.618)**
$\Delta \ln(E^H/E^M)_{o,t}$	-0.003	0.758	0.728	0.073	0.795	0.743	
	(0.288)	(0.124)***	(0.127)***	(0.167)	(0.103)***	(0.078)***	
$\Delta(w^H - w^M)^{LIS}_{o,t}$		-0.011	-0.018				
		(0.005)**	(0.005)***				
$\Delta Gini^{WIDER}_{o,t}$				-2.768			
				(2.035)			
$\Delta(\ln w^H - \ln w^M)^{LIS}_{o,t}$					-0.556	-0.577	
					(0.101)***	(0.109)***	
R-squared	-0.005	0.050	0.063	0.015	0.056	0.064	0.184
F-Statistics	73.484	45.295	33.628	60.657	54.719	37.190	72.462
Observations	3581	2766	2766	3407	2766	2766	3581
B. Dependent variable: Change in Ln(Middle Educated/Low Educated)							
$RSH_{j,t}$	-0.160	0.410	0.351	-0.057	0.394	0.353	-0.267
	(0.417)	(0.607)	(0.632)	(0.516)	(0.602)	(0.633)	(0.489)
$\Delta \ln(E^M/E^L)_{o,t}$	0.260	0.145	0.446	0.253	0.176	0.444	
	(0.122)**	(0.186)	(0.165)***	(0.141)*	(0.187)	(0.159)***	
$\Delta(w^M - w^L)^{LIS}_{o,t}$		0.064	-0.001				
		(0.054)	(0.038)				
$\Delta Gini^{WIDER}_{o,t}$				-3.570			
				(0.838)***			
$\Delta(\ln w^M - \ln w^L)^{LIS}_{o,t}$					0.075	0.033	
					(0.641)	(0.519)	
R^2	0.009	0.022	0.121	0.086	0.002	0.121	0.238
F-Statistics	76.745	25.830	45.230	63.064	53.952	50.376	74.811
Observations	3562	2497	2497	3351	2497	2497	3562
Origin country controls			√			√	
Decade × Orig. country FE							√

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors in parentheses are robust, clustered by CZ and origin country. Regressions include fixed effects for cantons, origin countries and decades, and are weighted using the total number of recent immigrants from origin country o in destination j at the beginning of the decade. $RSH_{j,t}$ is instrumented with $\widetilde{RSH}_{j,1970}$. $\Delta \ln(E^H/E^M)_{o,t}$ ($\Delta \ln(E^M/E^L)_{o,t}$) is the decennial change in log ratio of highly to middle educated (middle to low educated) population shares. $\Delta(w^H - w^M)^{LIS}_{o,t}$ is the change in the LIS wage difference (divided by 1000) between highly and middle educated (middle and low educated) workers in origin country o. $\Delta(\ln w^H - \ln w^M)^{LIS}_{o,t}$ is the change in log wage difference for highly and middle educated (middle and low educated) workers in origin country o, respectively. $\Delta Gini^{WIDER}_{o,t}$ represent the decennial change of the Gini index (using UNU WIDER data) in origin country o. Additional origin country controls are the change in real per capita GDP, the change in the polity IV measure and conflict dummies.

Table 5.A.6: Determinants of the Change in Skill Group Shares, 1980 - 2010, 2SLS, Excluding Cells with Zero Shares

	(1)	(2)	(3)	(4)	(5)
	A. Dependent variable: Change in Share of Highly Educated				
$RSH_{j,t}$	0.380	0.404	0.444	0.403	0.367
	(0.152)**	(0.159)**	(0.187)**	(0.159)**	(0.151)**
$\Delta EDUSH^{H}_{o,t}$	0.969	0.875	1.091	0.955	
	(0.313)***	(0.201)***	(0.249)***	(0.233)***	
$\Delta Gini^{WIDER}_{o,t}$		-0.373		-0.270	
		(0.204)*		(0.186)	
$\Delta GDPPC_{o,t}$		0.978	1.546	-0.752	
		(0.517)*	(0.347)***	(0.809)	
$\Delta Gini^{LIS}_{o,t}$			-0.526		
			(0.195)***		
$\Delta PolityIV_{o,t}$				-0.003	
				(0.003)	
$Conflict_{o,t}$				-0.139	
				(0.060)**	
R^2	0.048	0.091	0.145	0.118	0.316
F-Statistics	72.0	57.9	52.1	57.9	70.9
	B. Dependent variable: Change in Share of Middle Educated				
$RSH_{j,t}$	-0.266	-0.287	-0.297	-0.288	-0.281
	(0.123)**	(0.123)**	(0.149)**	(0.122)**	(0.115)**
Dedush2midBL	0.599	0.728	0.769	0.700	
	(0.210)***	(0.235)***	(0.246)***	(0.256)***	
$\Delta Gini^{WIDER}_{o,t}$		-0.052		-0.008	
		(0.177)		(0.237)	
$\Delta GDPPC_{o,t}$		1.798	1.670	1.618	
		(0.264)***	(0.248)***	(0.868)*	
$\Delta Gini^{LIS}_{o,t}$			0.075		
			(0.267)		
$\Delta PolityIV_{o,t}$				-0.002	
				(0.003)	
$Conflict_{o,t}$				0.005	
				(0.057)	
R^2	0.048	0.125	0.125	0.126	0.303
F-Statistics	72.0	58.0	52.1	58.0	70.9
	C. Dependent variable: Change in Share of Low Educated				
$RSH_{j,t}$	-0.112	-0.116	-0.145	-0.114	-0.086
	(0.047)**	(0.063)*	(0.078)*	(0.065)*	(0.060)
$\Delta EDUSH^{L}_{o,t}$	0.625	0.589	0.911	0.677	
	(0.397)	(0.294)**	(0.290)***	(0.277)**	
$\Delta Gini^{WIDER}_{o,t}$		0.482		0.358	
		(0.255)*		(0.275)	
$\Delta GDPPC_{o,t}$		-2.779	-3.312	-1.080	
		(0.543)***	(0.354)***	(1.120)	
$\Delta Gini^{LIS}_{o,t}$			0.447		
			(0.280)		
$\Delta PolityIV_{o,t}$				0.003	
				(0.002)	
$Conflict_{o,t}$				0.139	
				(0.088)	
R^2	0.028	0.204	0.294	0.227	0.446
F-Statistics	72.1	58.1	52.2	58.0	70.9
Observations	2904	2756	2200	2756	2904
Decade × Orig. country FE					√

Note: See the notes to Table 5.1. Zero Cells were excluded in case of the New Census (the year 2010) as the New Census is not based on a full inventory count of the population.

5.B Results Reproduced from Beerli and Peri (2015)

Beerli and Peri (2015) exploit the fact that the FMP policy was implemented gradually with two different schedules in two parts of Switzerland: the border region, which incorporates the area close to the national border of Switzerland, and the non-border region, the rest of the country. As discussed and documented in detail in their paper, the Federal Council announced in 1999 that access to the border region will be first completely liberalised for cross-border workers, who commuted to work from a neighbouring country work in the Swiss border region. Although this full liberalisation would be implemented only after 2004, cantonal immigration offices in the border region already anticipated the policy change and started to issue cross-border permits in a more liberal fashion after the announcement in 1999. After 2007, access for all types of immigrants (cross-border workers and new residing immigrants) was completely liberalised to both regions. Consequently, the authors estimate the effect of the announcement (*Phase 1*, 1999-2004) and the implementation (*Phase 2*, after 2004) of the FMP policy on the skill-mix of new immigrants with the following difference-in-difference design with separate regressions for each education group $e \in \{\text{high,middle,low}\}$:

$$\begin{aligned} EDUSH^e_{j,t} \;=\; & \alpha_j + \alpha_t + \beta^e_1 \left[BR_j \times 1(2000 \leq year < 2004) \right] \\ & + \beta^e_2 \left[BR_j \times 1(2004 \leq year \leq 2010) \right] + X'_{j,t}\gamma + \epsilon_{j,t} \end{aligned} \tag{5.11}$$

where $EDUSH^e_{j,t}$ is the share of new immigrant workers with education e on the total new immigrant employment in area j and year t. Areas are either municipalities or commuting zones as specified in Section 3. α_j and α_t absorb area and year fixed effects, respectively. BR_j is a dummy indicating whether an area belongs to the border region. $1(2000 \leq year < 2004)$ and $1(2004 \leq year \leq 2010)$ are dummies for Phase 1 and Phase 2 during which the two regions were differently open for new immigrants due to the implementation of the FMP policy.

The estimates for β^e_1 and β^e_2 indicate, whether the share of a skill group among new immigrants changed *differentially* in the border region (compared to the rest of the country) during the implementation of the FMP policy. Table 5.B.1 shows the estimates of specification (5.11) separately for municipalities (Columns 1 to 3) and Commuting Zones (Columns 4 to 6). Column 1 and 4 control for year fixed effects only whereas Column 2 and 5 also include area fixed effects. In Column 3 and 6 controls for education-specific demand trends in areas, $X_{j,t}$, are included (see Beerli and Peri (2015) for a discussion).[50] The estimate of BR_m show that the share of highly (middle) educated was about 5 (15) percentage points higher in the border region prior to the FMP policy whereas the share of low educated was 20 percentage points lower compared to the non-border region. The table shows further, that there was no effect on the skill-mix of immigrants of Phase 1, i.e. the announcement. During Phase 2, however, the share of middle educated decreased and the share of low educated increased among new immigrants in the border region compared to the non-border region. This indicates that the FMP policy lead to a larger (smaller) inflow of low (middle educated) educated.

For their analysis, Beerli and Peri (2015) use the Swiss Earnings Structure Survey (SESS) which is available biannually from 1994 to 2010 and only for workers employed in the private sector. The SESS defines new immigrants based on whether they have a short-term work permit (permit B and L) and explicitly includes cross-border workers (permit G). In contrast, in the Swiss

[50] In particular, column 3 and 6 include so-called Bartik (1991) measures that controls for demand trends at the local level. The basic idea is that industry-specific demand trends at the national level affect commuting zones differently to the degree they are specialised in some industries rather than others. Sector driven for education group $e \in \{H, M, L\}$ in a commuting zone j in year t as:

$$\widetilde{EMP}^{e}_{j,t} = \sum_{i \in \{1,50\}} \left(EMP^{e}_{i,j,1994} \times \frac{EMP^{e}_{-j,i,t}}{EMP^{e}_{-j,i,1994}} \right) \tag{5.12}$$

where $EMP^{e}_{i,j,1994}$ is the employment level of education group e in commuting zone j and (2-digit) industry i in the the year, 1994. $\frac{EMP^{e}_{-j,i,t}}{EMP^{e}_{-j,i,1994}}$ is the education group employment growth factor between 1994 and year t for the industry nationally, excluding the commuting zone of interest.

Table 5.B.1: Reprint of Table 5 from Beerli and Peri (2015): Effect of the FMP Policy on the Skill Composition of Immigrants

	(1)	(2)	(3)	(4)	(5)	(6)
A. Dependent variable: Share of highly educated among new immigrants						
$BR_j \cdot 1(2000 \leq year < 2004)$	0.006	-0.009	-0.008	0.002	-0.003	-0.001
	(0.018)	(0.014)	(0.015)	(0.020)	(0.021)	(0.021)
$BR_j \cdot 1(2004 \leq year \leq 2010)$	0.009	0.008	0.009	-0.003	-0.001	0.000
	(0.025)	(0.017)	(0.017)	(0.026)	(0.021)	(0.020)
BR_j	0.050			0.044		
	(0.025)**			(0.025)*		
R-squared	0.134	0.721	0.721	0.237	0.888	0.889
B. Dependent variable: Share of middle educated among new immigrants						
$BR_j \cdot 1(2000 \leq year < 2004)$	-0.000	-0.008	-0.011	0.004	0.008	0.004
	(0.017)	(0.018)	(0.019)	(0.018)	(0.018)	(0.018)
$BR_j \cdot 1(2004 \leq year \leq 2010)$	-0.118	-0.106	-0.110	-0.111	-0.104	-0.108
	(0.024)***	(0.026)***	(0.026)***	(0.024)***	(0.025)***	(0.024)***
BR_j	0.154			0.151		
	(0.032)***			(0.033)***		
R-squared	0.074	0.579	0.580	0.149	0.778	0.780
C. Dependent variable: Share of low educated among new immigrants						
$BR_j \cdot 1(2000 \leq year < 2004)$	-0.006	0.017	0.019	-0.005	-0.005	-0.003
	(0.024)	(0.018)	(0.018)	(0.026)	(0.029)	(0.029)
$BR_j \cdot 1(2004 \leq year \leq 2010)$	0.109	0.099	0.101	0.114	0.105	0.108
	(0.026)***	(0.025)***	(0.025)***	(0.027)***	(0.024)***	(0.025)***
BR_j	-0.204			-0.195		
	(0.047)***			(0.048)***		
R-squared	0.181	0.718	0.719	0.286	0.894	0.894
Observations	12253	12253	12248	946	946	943
Year Fixed Effects	√	√	√	√	√	√
Area Fixed Effects		√	√		√	√
Bartik			√			√

Note: ***, **, *, denote statistical significance at the 1%, 5% and 10% level, respectively. Robust standard errors, clustered by commuting zone, are given in parentheses. BR_m is one for municipalities (commuting zones) in the border region. $1(2000 \leq year < 2004)$ and $1(2004 \leq year \leq 2014)$ are dummies for the differential opening in *Phase 1*, from 1999 to 2004, and *Phase 2*, from 2004 to 2010, respectively. In each panel, the third and sixth column include the log of the education specific Barik control as specified in the text. Regressions are weighted using the total number of new immigrants of cells.

Census we define new immigrants as individuals (working and non-working) *residing* in Switzerland no more than 5 years. Thus, the difference in the effect of the FMP policy on the share of highly educated new immigrants (zero in Table 5.B.1 and negative in Table 5.3) may be attributed to some degree to the difference in the data used and to some degree to the difference in the identification strategy used. The consistent finding from both results is that the FMP policy increased the marginal benefit of immigration more for workers at the lower end of the skill distribution.

To be still closer in spirit to our baseline regression specification, we can extend equation (5.11) as follows. First, we include a full set of border region × year interactions to analyse whether the skill-mix of new immigrants changed differentially in the border region prior to the FMP policy additionally to the effects of the policy in Phase 1 and Phase 2. Second, we include the (instrumented) routine share of commuting zone's employment in 1990 interacted with a full set of year dummies:

$$\begin{aligned} EDUSH^e_{j,t} \;=\; & \alpha_j + \alpha_t + \sum_{t=1994}^{1996} \gamma^e_t \left[BR_j \times 1(year = t)\right] \\ & + \sum_{t=2000}^{2010} \beta^e_t \left[BR_j \times 1(year = t)\right] \\ & + \sum_{t=1996}^{2010} \delta^e_t \left[RSH_{j,1990} \times 1(year = t)\right] + \epsilon_{j,t} \end{aligned} \tag{5.13}$$

In this specification, γ^e_t and β^e_t measure the differential change of an education group share of employment among new immigrants prior and after the implementation of the FMP policy, respectively.[51] The estimates of the coefficient δ^e_t indicate the effect of routinisation on the share of an education group relative to the initial year 1994.

Figure 5.B.1 plots γ^e_t and β^e_t with the share of highly, middle and low educated among new immigrants as the dependent variable in Panel A, B and C, respectively. This confirms the results above, i.e. the share of middle (low) educated decreased (increased) in the border region after the introduction of the FMP policy. In addition, the figure shows that the skill-mix of new immigrants did not change differentially across the two regions prior to the FMP policy.[52]

[51] Note that these effects are relative to the last year prior to the policy onset 1998.

[52] The estimates of the coefficients δ^e_{2010} have the expected sign for each education group, i.e. $\delta^H_{2010} = 0.19$, $\delta^M_{2010} = -0.94$ and $\delta^L_{2010} = 0.75$. These results are available from the authors on request.

Figure 5.B.1: Effect of the FMP Policy on the Skill Composition of Immigrants, Event Analysis

Note: The solid line depicts an estimate of γ_t^e and β_t^e for highly, middle and low educated among new immigrants as specified in equation (5.13) using commuting zones as areas. The dashed lines mark the 95% confidence intervals of each estimate. Standard errors are clustered at the commuting zone level.

5.C Data Construction

5.C.1 Measuring Employment

We measure labour supply in full time equivalents (FTEs) based on weekly hours worked. For the 1970, 1980, 1990 and 2000 censuses, however, only the following 4 categories were available for weekly hours worked instead of exact hours. We therefore used the 1991 Swiss Labour Survey (Schweizerische Arbeitskräfteerhebung, SAKE) which contains the same 4 categories but also the actual hours worked in order to arrive at the following weighting scheme:

- 1-5 hours per week: 0.05 FTEs
- 6-24 hours per week: 0.35 FTEs
- 25-41 hours per week: 0.7 FTEs
- 42 or more hours per week: 1 FTEs

In the 2010 census on the other hand, weekly hours were exactly measured. We divided them by 42 which corresponds to a full time working week measured in hours. Hours above 42 were capped, i.e. an individual working more than 42 hours per week was counted as one full time equivalent. Since the new census is conducted as a representative sample, individuals were then weighted according to the official weights contained in the data.

5.C.2 Defining Occupations

The Swiss Census contains two classifications of occupations for all survey years. The ISCO-88-COM is a version of the ISCO-88 and internationally comparable in principle. The Swiss Nomenclature of Occupations 2000 (SNO-2000) is a national classification scheme not comparable to those of other countries. However, it is straightforward to map it to the ISCO-88 or to US occupation classifications.

In case of the ISCO-88-COM, about a third of the two-digit ISCO-classes contain no observations prior to 2010 and about half of two-digit classes are missing prior to 1990. This makes it impossible to keep track of the two-digit-classes in detail. Furthermore, some of the occupations show an unplausibly volatile development in recent years, in particular ISCO no. 93, which jumps from less than 90'000 in 1980 to over 430'000 in 1990 and then falls again to less than 20'000 in 2000. According to the Statistical Office of Switzerland, ISCO no. 93 contains a vast number of employees (especially in 1990) which could not be appropriately allocated to the ISCO-classification and thus were assigned to the broad class ISCO no. 93. However, one-digit ISCO-classes seem to yield plausible results and can be used to check our descriptive results (where ISCO 93 must be excluded from the analysis).

Contrary to the ISCO-88-COM, the SBN-2000 has entries in almost all of the two-digit classes for every survey year. Due to this fairly complete picture and the existence of reliable and complete keys, we are able to map the SNO-2000 occupations into the ISCO-88-COM by two steps. First, we mapped the SNO-2000 to the older SNO-1990 [53] and then, using another crosswalk and thereby following Basten and Siegenthaler (2013), mapped

[53] In principle, the crosswalk SNO-2000 / SNO-1990 only allows for a complete matching from SNO-1990 to SNO-2000 occupations but not the other way round. However, since the mapping in almost all cases is a 1-to-1 mapping, almost no occupations get lost (i.e. classified missing) if one maps the SNO-2000 to SNO-1990.

those classes into the ISCO-88-COM. This procedure resulted in 26 two-digit ISCO-88-COM classes with no missing entries and, in particular, implausibly volatile occupation classes do not occur anymore. However, similar to the ISCO-classification, the SNO-2000 contains one extraordinary volatile class, namely SNO-2000 no. 93, whose inclusion would yield to seriously distorded labour-shares for the other occupations. SNO 93 along with SNO 92 explicitly contain occupations which could not be classified by the statistical agency. Hence we excluded SNO 92 along with SNO 93 from the analysis. Albeit we have to take into account that the constructed ISCO-88-COM classes may introduce some inaccuracy, it gives us the possibility to work with about 26 ISCO-classes instead of only 1-digit-ISCOs contained in the censuses. The resulting occupations exhibit a plausible development over the years and the results obtained in our paper fit nicely into the results found by the literature for other countries.

5.C.3 Gini Coefficients and Education Specific Wages in Origin Countries

Our main datasource for gini coefficients is the UNU-WIDER World Income Inequality Database, Version 2.0c, May 2008. Generally, we include only inequality measures based on disposable income. Due to the numerous amount of studies in the case of some countries, we computed averages of the gini coefficients at hand. The averaging resulted in little differences compared to relying only on a single datasource, if available for any given country, but has the advantage that the resulting time series extend over a much longer time horizon. For other countries such as China or India, inequality measures based on disposable income were scant or unavailable. In these cases we also included studies that calculated gini coefficients based on consumption or net income data.

As a main alternative, we use gini coefficients and education specific wages from the Luxembourg Income Study's (LIS) 'Key Figures' (Version 3), which is only available for a subset of origin countries and years. Since a number of comparison issues arise when working with educational attainment information directly using the LIS, we follow GH and use the quantiles of a country's earnings distribution to gauge wages of different education groups. The LIS provides information on the ratio of the the 90th percentile to the 10th percentile and of the 90th percentile to the median for various countries earnings distributions in different years. We linearly interpolate the ratios in missing years between available waves and extrapolate trends up to 10 years to minimize the loss of observations.[54] We approximate median income by origin country with GDP per capita from Heston, Summers and Aten (2011) and use the ratios from the LIS to gauge incomes for the 90th and 10th percentile. We use the median wage as our wage measure for middle educated workers, and the 90th and 10th percentile as a wage measure for highly and low educated, respectively.

5.C.4 Offshorability

We obtain offshorability measures from two sources. Using several hundred cases of offshoring in Europe, Goos, Manning and Salomons (2011) construct an index of how offshorable an occupation is (see Table 4 in Goos, Manning and Salomons, 2011). As they work with two-digit-ISCOs, we can directly map those indices to our dataset. Blinder and Krueger (2013) on the other hand, report survey measurements of offshorability for US occupations, industries and various personal characteristics of their dataset such as offshorability by education level. Matching the latter to our dataset proved again straightforward as similar educational attainment measures were contained in the Swiss census. From these offshorability measures by occupations and educational levels, we compute offshorability indices for each commuting zone in Switzerland in the following ways:

$$OFFSH^{occ}_{j,t} = \sum_k \frac{L_{j,t,k}}{L_{j,t}} OFFSH_{t,k}$$

[54] Even though we impute the data as described, data for most developed countries in our sample (China, Chile, Algeria, India, Iran, Tunisia, Turkey and Vietnam) remains missing. For the group of Ex-Yugoslavian countries, only data from Slovenia is available. Thus, we used then the Gini of Slovenia to gauge the skill returns in all Ex-Yugoslavian countries.

$$OFFSH^{skill}_{j,t} = \sum_e \frac{L_{j,t,e}}{L_{j,t}} OFFSH_{t,e}$$

where $OFFSH_{t,k}$ is the offshorability measures for occupations from Goos, Manning and Salomons (2011). And $OFFSH_{t,e}$ are the offshorability measures for skill groups *high*, *middle* and *low*, computed by using the indices provided by Blinder and Krueger (2013) for 6 different educational attainments.

5.D Additional Theory Appendix

Much of the existing literature on international migration flows (e.g. Borjas, 1999) assumes the selection of migrants depends on the *relative* returns to skills rather than absolute wage differences as posited by GH. To compare their framework to the traditional model, GH derive the log-odds ratio of migration and the corresponding sorting equation based on a adaption of their model with log-utility. Thus, the prospective utility of a migrant i with education $e \in \{H, M, L\}$ from origin country o in destination j is

$$U^e_{i,o,j} = \left(W^e_{i,j} - C^e_{i,o,j}\right)^\lambda \exp\left(\mu^e_{i,o,j}\right) \tag{5.14}$$

where $\lambda > 0$ and $\mu^e_{i,o,j}$ follows an i.i.d. extreme value distribution. Then, the her log-odds ratio of migration is given by

$$\ln \frac{L^e_{o,j}}{L^e_o} = \lambda \left(\ln W^e_j - \ln W^e_o\right) - \lambda m^e_{o,j} \tag{5.15}$$

where $m^e_{o,j} = \left(f_{o,j} - g^e_{o,j}\right) / W^e_j$ are education-specific migration costs proportional to the wage in destination j. Then, the sorting equation for highly educated relative to middle educated migrants can be expressed as

$$\ln \frac{L^H_{j,o}}{L^M_{j,o}} = \lambda \left(\delta^H_j - \delta^H_o\right) - \lambda \left(m^H_{o,j} - m^M_{o,j}\right) + \ln \frac{L^H_o}{L^M_o} \tag{5.16}$$

where $\delta^H = \ln W^H - \ln W^M$ is the return to skill for highly educated over middle educated.

5.E More Details on Immigration Policies 1980-2010

Since the early 1970s, Swiss immigration policy was characterised by global quotas for different types of residency permits. At the end of each year, the federal government determined the number of new permits for different residency categories which would be allocated to Cantonal administrations proportional to population size. Cantonal administrations, in turn, would grant residency permits to firms which wanted to hire foreign workers if no equally qualified natives could be found for a given job (the so-called "priority requirement").[55]

The first change to this policy came in 1991, when the federal government introduced new regulations requiring residency permits to be granted primarily to immigrants from EU17/EFTA countries while restricting immigration from other countries to 'highly qualified' individuals (Bundesrat, 1991).[56] Global quotas remained in place, however, until 2002.

After rejecting the proposal to join the European Economic Area in 1992 in national referendum, Switzerland and the EU signed a package of bilateral agreements on June 21, 1999, including the free movement policy for EU workers.[57] Accepting the free movement policy, one of the central pillars of the EU's 'acquis', was a necessary concession of Switzerland to achieve liberalised trade with the EU. Details about the liberalization process were publicly announced by the federal administration (Bundesrat, 1999) and in 2000, the entire bilateral package was approved by a referendum in Switzerland in May 2000.

[55] Exempted from the quota were newly immigrating family members of foreign residents and previous seasonal workers who spent at least five consecutive seasons working in Switzerland (Sheldon, 2007).

[56] 'High qualification' mainly meant "managers, specialists and qualified professionals" but was deliberately defined broadly to include also investors, workers who specialised in the construction of certain facilities or tunnels or workers with special knowledge (Bundesrat, 1991, 2002). Third party countries included workers from eastern European countries as well as Bulgaria and Romania.

[57] The package of bilateral agreements included also agreements on the reduction of technical barriers to trade, the liberalisation of trade with agricultural good and public procurement, transport and the participation of Switzerland in the EU?s research framework programmes.

The enactment of the Free Movement of Persons Agreement (AFMP) with the European Union (EU) on June 1 2002 marked the second and most important change in the Swiss immigration policy. With the AFMP, the federal government introduced separate quotas for immigrant workers from EU17/EFTA and Non-EU origin countries. After June 1 2004 the the priority requirement for natives was abolished and the border region was completely liberalised for cross-border workers from EU17/EFTA countries but not for resident immigrants.[58] From June 1 2007, workers from EU17/EFTA countries had unrestricted access to the Swiss labor market.

For Eastern European (EU8) countries, who jointed the EU in 2004, the AFMP was enacted on June 1, 2006, with the introduction of a separate quota. Prior to this date, they were part of the group of Non-EU countries facing similar qualification restriction. On May 1st, 2011, quotas for EU8 countries were abolished granting free movement.[59]

Access was facilitated in a similar step-wise fashion for Romania and Bulgaria starting on June 1, 2009, by granting them a separate quota and the abolition of qualification restrictions. These quotas will be in place until mid 2016.

Access for immigrants from Non-EU countries kept being subject to quotas and similar qualification restrictions as in the 1990s through our analysis period.(Bundesrat, 2002).[60] The quota was separated from European countries on June 1, 2002.

[58] EU17 countries include Germany, France, Italy, Luxembourg, Netherlands, Denmark, Ireland, United Kingdom, Greece, Portugal, Spain, Finland, Austria, Sweden, Cyprus and Malta. EFTA countries include Iceland, Norway and Liechtenstein. Cross-border workers are individuals commuting to work in Switzerland from one of the neighbouring countries. Prior to June 1 2004, these workers were also subject entry restrictions. As cross-border workers are not in our data set, we refer to Beerli and Peri (2015) for an extensive discussion.

[59] Until April 2014, the Federal Council had the opportunity to unilaterally re-introduce temporary quotas for immigrant workers from the EU, if immigration reached certain levels, through a so-called 'Protection Valve'. Quotas were re-introduced between between April 2012 and 2014 for workers from EU8 countries and between April 2013 and 2014 for EU17/EFTA countries.

[60] Even though early reports from the federal council, (Bundesrat, 1991), are less precise and detailed about qualification requirements for Non-EU immigrants than later reports, e.g. (Bundesrat, 2002), officials from the State Secretariat for Migration assured that these requirements were very similar.

Bibliography

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2012. "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration." *American Economic Review*, 102(5): 1832–1856.

Acemoglu, Daron, and David Autor. 2011. "Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*. Vol. 4, Part B, , ed. David Card and Orley Ashenfelter, 1043 – 1171. Elsevier.

Amstad, Marlene, and Andreas M. Fischer. 2009. "Are Weekly Inflation Forecasts Informative?" *Oxford Bulletin of Economics and Statistics*, 71(2): 237–252.

Anderson, Richard G. 2006. "Replicability, Real-Time Data, and the Science of Economic Research: FRED, ALFRED, and VDC." *Federal Reserve Bank of St. Louis Review*, 88(1): 81–93.

Aruoba, S. Borağan. 2008. "Data Revisions Are Not Well Behaved." *Journal of Money, Credit and Banking*, 40(2-3): 319–340.

Autor, David, and David Dorn. 2009. "This Job Is ''Getting Old'': Measuring Changes in Job Opportunities Using Occupational Age Structure." *American Economic Review*, 99(2): 45–51.

Autor, David H., and David Dorn. 2013. "The Growth of Low Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review*, 5(103): 1553–1597.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics*, 118(4): 1279–1333.

Barro, Robert J., and Jong Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950 to 2010." *Journal of Development Economics*, 104(0): 184 – 198.

Barsky, Robert B., and Jeffrey A. Miron. 1989. "The Seasonal Cycle and the Business Cycle." *Journal of Political Economy*, 97(3): 503–534.

Bartel, Ann P. 1989. "Where Do the New US Immigrants Live?" *Journal of Labor Economics*, 7(4): 371–391.

Bartik, Timothy J. 1991. "Who Benefits from State and Local Economic Development Policies?" *Books from Upjohn Press.*

Basten, Christoph, and Michael Siegenthaler. 2013. "Do Immigrants Take or Create Residents' Jobs?" *KOF Working Papers*, 335.

Beaulieu, J. Joseph, and Jeffrey A. Miron. 1992. "A Cross Country Comparison of Seasonal Cycles and Business Cycles." *The Economic Journal*, 102(413): 772–778.

Beaulieu, J. Joseph, and Jeffrey A. Miron. 1993. "Seasonal Unit Roots in Aggregate US Data." *Journal of Econometrics*, 55(1-2): 305–328.

Beerli, Andreas, and Giovanni Peri. 2015. "The Labor Market Effects of Opening the Border: New Evidence from Switzerland." *NBER Working Paper*, , (21319).

Bell, William R., and Steven C. Hillmer. 1984. "Issues Involved with the Seasonal Adjustment of Economic Time Series." *Journal of Business & Economic Statistics*, 2(4): 291–320.

Belot, Michèle V.K., and Timothy J. Hatton. 2012. "Immigrant Selection in the OECD." *The Scandinavian Journal of Economics*, 114(4): 1105–1128.

Bernhard, Severin. 2016. "A Real-Time GDP Data Set for Switzerland." *SNB Economic Studies*, 9/2016.

Bertoli, Simone, Herbert Brücker, Giovanni Facchini, Anna Maria Mayda, and Giovanni Peri. 2012. "Understanding Highly Skilled Migration in Developed Countries: The Upcoming Battle for Brains." In *Brain Drain and Brain Gain. The Global Competition to Attract High-Skilled Migrants.* , ed. Tito Boeri, Herbert Brücker, Frédéric Docquier and Hillel Rapoport. Oxford University Press.

Blinder, Alan S., and Alan B. Krueger. 2013. "Alternative Measures of Offshorability: A Survey Approach." *Journal of Labor Economics*, 31(2): 97–128.

Boldin, Michael, and Jonathan H. Wright. 2015. "Weather Adjusting Economic Data." *Brookings Papers on Economic Activity*, , (2): 227–278.

Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." *American Economic Review*, 77(4): 531–553.

Borjas, George J. 1999. "Chapter 28 - The Economic Analysis of Immigration." In *Handbook of Labor Economics*. Vol. 3, Part A, , ed. Orley C. Ashenfelter and David Card, 1697 – 1760. Elsevier.

Borjas, George J. 2008. "Labor Outflows and Labor Inflows in Puerto Rico." *Journal of Human Capital*, 2(1): 32–68.

Boschan, Charlotte, and Gerhard Bry. 1971. *Cyclical Analysis of Time Series: Procedures and Computer Programs.* NBER.

Brücker, H., S. Capuano, and A. Marfouk. 2013. "Education, Gender and International Migration: Insights from a Panel-Dataset 1980-2010." mimeo.

Bundesrat. 1991. "Bericht des Bundesrates zur Auslaender- und Fluechtlingspolitik." *http://www.amtsdruckschriften.bar.admin.ch/viewOrigDoc.do?id=10051898*.

Bundesrat. 1999. "Botschaft zur Genehmigung der sektoriellen Abkommen zwischen der Schweiz und der EG vom 23. Juni 1999." *https://www.admin.ch/opc/de/federal-gazette/1999/6128.pdf*.

Bundesrat. 2002. "Botschaft zum Bundesgesetz über die Ausländerinnen und Ausländer." *https://www.admin.ch/opc/de/federal-gazette/2002/3709.pdf*.

Burman, Peter J. 1980. "Seasonal Adjustment by Signal Extraction." *Journal of the Royal Statistical Society. Series A (General)*, 321–337.

Burridge, Peter, and Kenneth F. Wallis. 1984. "Unobserved-Components Models for Seasonal Adjustment Filters." *Journal of Business & Economic Statistics*, 350–359.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference With Multiway Clustering." *Journal of Business & Economic Statistics*, 29(2): 238–249.

Canova, Fabio, and Bruce E. Hansen. 1995. "Are Seasonal Patterns Constant over Time? A Test for Seasonal Stability." *Journal of Business & Economic Statistics*, 13(3): 237–252.

Canova, Fabio, and Eric Ghysels. 1994. "Changes in Seasonal Patterns: are they Cyclical?" *Journal of Economic Dynamics and Control*, 18(6): 1143–1171.

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics*, 19(1): 22–64.

Castle, Jennifer L., and Colin Ellis. 2002. "Building a Real-Time Database for GDP(E)." *Bank of England Quarterly Bulletin*, 42(1): 42–49.

Chaloff, Jonathan, and Georges Lemaitre. 2009. "Managing Highly-Skilled Labour Migration: A Comparative Analysis of Migration Policies and Challenges in OECD Countries." *OECD Social, Employment and Migration Working Papers*, , (79).

Chiquiar, Daniel, and Gordon H. Hanson. 2005. "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States." *Journal of Political Economy*, 113(2): 239–281.

Chow, Gregory C., and A. Lin. 1971. "Best linear unbiased interpolation, distribution, and extrapolation of time series by related series." *The Review of Economics and Statistics*, 53(4): 372–375.

Cimadomo, Jacopo. 2016. "Real-Time Data and Fiscal Policy Analysis: A Survey of the Literature." *Journal of Economic Surveys*, 30(2): 302–326.

Clark, Ximena, Timothy J. Hatton, and Jeffrey G. Williamson. 2007. "Explaining US Immigration, 1971-1998." *The Review of Economics and Statistics*, 89(2): 359–373.

Cleveland, William P. 1972. *Analysis and Forecasting of Seasonal Time Series.* PhD Dissertation, University of Wisconsin.

Corradi, Valentina, Andres Fernandez, and Norman R. Swanson. 2009. "Information in the Revision Process of Real-Time Datasets." *Journal of Business & Economic Statistics*, 27(4): 455–467.

Croushore, Dean. 2011*a*. "Forecasting with real-time data vintages." *The Oxford Handbook of Economic Forecasting*, 247–267.

Croushore, Dean. 2011*b*. "Frontiers of Real-Time Data Analysis." *Journal of Economic Literature*, 49(1): 72–100.

Croushore, Dean, and Charles L. Evans. 2006. "Data Revisions and the Identification of Monetary Policy Shocks." *Journal of Monetary Economics*, 53(6): 1135–1160.

Croushore, Dean, and Tom Stark. 2001. "A Real-Time Data Set for Macroeconomists." *Journal of Econometrics*, 105(1): 111–130.

Croushore, Dean, and Tom Stark. 2003. "A Real-Time Data Set for Macroeconomists: Does the Data Vintage Matter?" *Review of Economics and Statistics*, 85(3): 605–617.

Cuche-Curti, Nicolas, Pamela Hall, and Attilio Zanetti. 2008. "Swiss GDP Revisions: A Monetary Policy Perspective." *Journal of Business Cycle Measurement and Analysis*, 2: 183–213.

Cunningham, Alastair, Jana Eklund, Chris Jeffery, George Kapetanios, and Vincent Labhard. 2012. "A State Space Approach to Extracting the Signal From Uncertain Data." *Journal of Business & Economic Statistics*, 30(2): 173–180.

Dagum, Estela B. 1980. *The X-11-ARIMA Seasonal Adjustment Method.* Statistics Canada.

Dagum, Estela B. 1988. *The X-11-ARIMA/88 Seasonal Adjustment Method.* Statistics Canada.

Dagum, Estela B., and Marietta Morry. 1984. "Basic Issues on the Seasonal Adjustment of the Canadian Consumer Price Index." *Journal of Business & Economic Statistics*, 2(3): 250–259.

Dagum, Estela B., and Normand Laniel. 1987. "Revisions of Trend-Cycle Estimators of Moving Average Seasonal Adjustment Methods." *Journal of Business & Economic Statistics*, 5(2): 177–189.

Depoutot, Raoul, and Christophe Planas. 2002. "Controlling Revisions in Arima-Model-Based Seasonal Adjustment." *Journal of Time Series Analysis*, 23(2): 193–213.

Dickey, David A, and Wayne A. Fuller. 1979. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American statistical association*, 74(366a): 427–431.

Dickey, David A., David P. Hasza, and Wayne A. Fuller. 1984. "Testing for Unit Roots in Seasonal Time Series." *Journal of the American Statistical Association*, 79(386): 355–367.

Diebold, Francis X., and Glenn D. Rudebusch. 1991. "Forecasting Output With the Composite Leading Index: A Real-Time Analysis." *Journal of the American Statistical Association*, 86(415): 603–610.

Dustmann, Christian, and Tommaso Frattini. 2014. "The Fiscal Effects of Immigration to the UK." *The Economic Journal*, 124(580): 593–643.

Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg. 2009. "Revisiting the German Wage Structure." *Quarterly Journal of Economics*, 124(2): 843–881.

Dustmann, Christian, Tommaso Frattini, and Ian P. Preston. 2013. "The Effect of Immigration Along the Distribution of Wages." *Review of Economic Studies*, 80(1): 145–173.

Economist. 2016. "The Trouble With GDP." *The Economist*, 28.04.2016.

Egginton, Don M., Andreas Pick, and Shaun P. Vahey. 2002. "Keep it Real!: A Real-Time UK Macro Data Set." *Economics Letters*, 77(1): 15e–20.

Eurostat. 2008*a*. "Eurostatistics. Data for Short-Term Economic Analysis." No. 6, Eurostat.

Eurostat. 2008*b*. "Eurostatistics. Data for Short-Term Economic Analysis." No. 8, Eurostat.

Eurostat. 2013. "Handbook on Quarterly National Accounts." Eurostat.

Faust, Jon, John H. Rogers, and Jonathan H. Wright. 2005. "News and Noise in G-7 GDP Announcements." *Journal of Money, Credit and Banking*, 37(3): 403–419.

Fernandez, Adriana Z., Houston Branch, Evan F. Koenig, and Alex Nikolsko-Rzhevskyy. 2012. "A Real-Time Historical Database for the OECD." *Federal Reserve Bank of Dallas, Globalization and Monetary Policy Institute Working Paper*, , (96).

Findley, David F., Brian C. Monsell, William R. Bell, Mark C. Otto, and Bor-Chung Chen. 1998. "New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program." *Journal of Business & Economic Statistics*, 16(2): 127–152.

Fischer, Björn, and Christophe Planas. 2000. "Large Scale Fitting of Regression Models with ARIMA Errors." *Journal of Official Statistics*, 16(2): 173–184.

Garratt, Anthony, and Shaun P. Vahey. 2006. "UK Real-Time Macro Data Characteristics." *The Economic Journal*, 116(509): F119–F135.

Ghysels, Eric, and Denise R. Osborn. 2001. *The Econometric Analysis of Seasonal Time Series.* Cambridge University Press.

Giannone, Domenico, Jérôme Henry, Magdalena Lalik, and Michele Modugno. 2012. "An Area-Wide Real-Time Database for the Euro Area." *Review of Economics and Statistics*, 94(4): 1000–1013.

Glejser, H., and P. Schavey. 1974. "An Analysis of Revisions of National Accounts Data for 40 Countries." *Review of Income and Wealth*, 20(3): 317–332.

Gómez, Víctor. 1999. "Three Equivalent Methods for Filtering Finite Nonstationary Time Series." *Journal of Business & Economic Statistics*, 17(1): 109–116.

Gomez, Victor, and Agustin Maravall. 1996. "Programs TRAMO and SEATS. Instructions for the User." *Documento de Trabajo, Banca de Espagna*, 9628.

Gomez, Victor, and Agustin Maravall. 2001. "Seasonal Adjustment and Signal Extraction Time Series." In *A Course in Time Series Analysis.* , ed. Daniel Pena, George C. Tiao and Ruey S. Tsay. John Wiley and Sons.

Goos, Maarten, Alan Manning, and Anna Salomons. 2011. "Explaining Job Polarization: The Roles of Technology, Offshoring and Institutions." Center for Economic Studies Discussions Paper Series 11.34.

Goos, Maarten, Alan Manning, and Anna Salomons. 2014. "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring." *American Economic Review*, 104(8): 2509–26.

Goos, Maarten, and Alan Manning. 2007. "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain." *Review of Economics and Statistics*, 89(1): 118–133.

Gould, Eric D., and Omer Moav. 2016. "Does High Inequality Attract High Skilled Immigrants?" *The Economic Journal*, 126(593): 1055–1091.

Granger, Clive W.J. 1979. "Seasonality: Causation, Interpretation, and Implications." In *Seasonal Analysis of Economic Time Series.* , ed. Arnold Zellner, 33–56. NBER.

Grogger, Jeffrey, and Gordon H. Hanson. 2011. "Income Maximization and the Selection and Sorting of International Migrants." *Journal of Development Economics*, 95(1): 42–57.

Grogger, Jeffrey, and Gordon H. Hanson. 2015. "Attracting Talent: Location Choices of Foreign-Born PhDs in the United States." *Journal of Labor Economics*, 33(S1): 5–38.

Harding, Don, and Adrian Pagan. 2002. "Dissecting the Cycle: A Methodological Investigation." *Journal of Monetary Economics*, 49(2): 365–381.

Heston, Alan, Robert Summers, and Bettina Aten. 2011. "Penn World Table Version 7.1." Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hillmer, Steven C., and George C. Tiao. 1982. "An ARIMA-Model-Based Approach to Seasonal Adjustment." *Journal of the American Statistical Association*, 77(377): 63–70.

Howrey, Philip E. 1978. "The Use of Preliminary Data in Econometric Forecasting." *The Review of Economics and Statistics*, 60: 193–200.

Huber, Peter, and Julia Bock-Schappelwein. 2014. "The Effects of Liberalizing Migration on Permanent Migrants' Education Structure." *Journal of Common Market Studies*, 52(2): 268–284.

Hylleberg, Svend, Robert F. Engle, Clive W.J. Granger, and Byung Sam Yoo. 1990. "Seasonal Integration and Cointegration." *Journal of Econometrics*, 44(1): 215–238.

Indergand, Ronald, and Stefan Leist. 2014. "A Real-Time Data Set for Switzerland." *Swiss Journal of Economics and Statistics*, 150(IV): 331–352.

Jacobs, Jan P.A.M., and Jan-Egbert Sturm. 2008. "The Information Content of KOF Indicators on Swiss Current Account Data Revisions." *Journal of Business Cycle Measurement and Analysis*, 4(2): 163–183.

Jacobs, Jan P.A.M., and Simon Van Norden. 2011. "Modeling Data Revisions: Measurement Error and Dynamics of "True" Values." *Journal of Econometrics*, 161(2): 101–109.

Johnson, Simon, William Larson, Chris Papageorgiou, and Arvind Subramanian. 2013. "Is Newer Better? Penn World Table Revisions and their Impact on Growth Estimates." *Journal of Monetary Economics*, 60(2): 255–274.

Jordan, Thomas J., Peter Kugler, Carlos Lenz, and Marcel R. Savioz. 2005. "GDP Data Revisions and Forward-Looking Monetary Policy in Switzerland." *The North American Journal of Economics and Finance*, 16(3): 351–372.

Kavajecz, Kenneth, and Sean Collins. 1995. "Rationality of Preliminary Money Stock Estimates." *The Review of Economics and Statistics*, 32–41.

Keane, Michael P., and David E. Runkle. 1990. "Testing the Rationality of Price Forecasts: New Evidence from Panel Data." *American Economic Review*, 80(4): 714–735.

Kerr, William R., and William F. Lincoln. 2010. "The Supply Side of Innovation: H-1B Visa Reforms and US Ethnic Invention." *Journal of Labor Economics*, 28(3): 473–508.

Kholodilin, Konstantin A., and Boriss Siliverstovs. 2012. "Assessing the Real-Time Informational Content of Macroeconomic Data Releases for Now-/Forecasting GDP: Evidence for Switzerland." *Jahrbucher fur Nationalokonomie und Statistik*, 232(4): 429–444.

Knetsch, Thomas A. 2009. "The German Real-Time Database Gerda: Content, Structure, Search and Download." Deutsche Bundesbank.

Koopman, S.J., A.C. Harvey, J.A. Doornik, and N. Shephard. 1996. *Stamp: Structural Time Series Analyser, Modeller and Predictor.* Chapman and Hall, London.

Kuznets, Simon. 1937. "National Income, 1929-1932." *NBER Bulletin*, 49.

Kwiatkowski, Denis, Peter C.B. Phillips, Peter Schmidt, and Yongcheol Shin. 1992. "Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How sure are we that Economic Time Series have a Unit Root?" *Journal of Econometrics*, 54(1-3): 159–178.

Mankiw, N. Gregory, and Matthew D. Shapiro. 1986. "News or Noise? An Analysis of GNP Revisions." *NBER Working Paper*.

Mankiw, N. Gregory, David E. Runkle, and Matthew D. Shapiro. 1984. "Are Preliminary Announcements of the Money Stock Rational Forecasts?" *Journal of Monetary Economics*, 14(1): 15–27.

Manski, Charles F. 2015. "Communicating Uncertainty in Official Economic Statistics: An Appraisal Fifty Years after Morgenstern." *Journal of Economic Literature*, 53(3): 631–653.

Maravall, Agustin. 1980. "Effects of Alternative Seasonal Adjustment Procedures on Monetary Policy." *Journal of Econometrics*, 14(1): 115–136.

Maravall, Agustin. 1985. "On Structural Time Series Models and the Characterization of Components." *Journal of Business & Economic Statistics*, 3(4): 350–355.

Marshall, Monty G., Ted Robert Gurr, and Keith Jaggers. 2014. "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2013." University of Maryland.

Mayda, Anna Maria. 2010. "International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows." *Journal of Population Economics*, 23(4): 1249–1274.

McElroy, Tucker. 2008. "Matrix Formulas for Nonstationary ARIMA Signal Extraction." *Econometric Theory*, 24(04): 988–1009.

McKenzie, David, and Hillel Rapoport. 2010. "Self-Selection Patterns in Mexico-US Migration: The Role of Migration Networks." *Review of Economics and Statistics*, 92(4): 811–821.

Michaels, Guy, Ashwini Natraj, and John Van Reenen. 2014. "Has ICT polarized Skill Demand? Evidence from Eleven Countries Over 25 Years." *Review of Economics and Statistics*, 96(1): 60–77.

Mincer, Jacob A., and Victor Zarnowitz. 1969. "The evaluation of economic Forecasts." In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance.* , ed. Jacob A Mincer, 3–46. NBER.

Miron, Jeffrey A. 1996. *The Economics of Seasonal Cycles.* Cambridge, Massachusetts:The MIT Press.

Monsell, Brian C. 2007. "The X-13A-S Seasonal Adjustment Program."

Morgenstern, Oskar. 1950. *On the Accuracy of Economic Observations.* . 1 ed., Princeton University Press.

Müller, Jürg. 2015. "Ökonomische Ratespiele." *Neue Zürcher Zeitung*, 16.05.2015.

Müller, Tobias, Noé Asensio, and Roman Graf. 2013. "Les effets de la libre circulation des personnes sur les salaires en Suisse." University of Geneva.

Newey, Whitney K., and Kenneth D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55: 703–708.

OECD. 2014. "International Migration Database." *http://www.oecd.org/els/mig/keystat.htm*.

OECD, Short-Term Economic Statistics Expert Group. 2002. "Harmonizing Seasonal Adjustment Methods in European Union and OECD Countries." *Note by the Secretariat.*

Oesch, Daniel, and Jorge Rodríguez Menés. 2011. "Upgrading or Polarization? Occupational change in Britain, Germany, Spain and Switzerland, 1990–2008." *Socio-Economic Review*, 9(3): 503–531.

Orphanides, Athanasios. 2001. "Monetary Policy Rules Based on Real-Time Data." *American Economic Review*, 964–985.

Orphanides, Athanasios, and Simon Van Norden. 2005. "The Reliability of Inflation Forecasts based on Output Gap Estimates in Real Time." *Journal of Money, Credit, and Banking*, 37(3): 583–601.

Ortega, Francesc, and Giovanni Peri. 2014. "Openness and Income: The Roles of Trade and Migration." *Journal of International Economics*, 92(2): 231–251.

Osborn, Denise R. 1990. "A Survey of Seasonality in UK Macroeconomic Variables." *International Journal of Forecasting*, 6(3): 327–336.

Pierce, David A. 1980. "Data Revisions with Moving Average Seasonal Adjustment Procedures." *Journal of Econometrics*, 14(1): 95–114.

Plickert, Philip. 2014. "Neue Statistik lässt Staatsschuldenquoten sinken." *Frankfurter Allgemeine Zeitung*, 21.01.2014.

Sax, Christoph. 2014. "Seasonal: R interface to X-13ARIMA-SEATS." R package version 0.60.0.

Sax, Christoph. 2015. "Seasonal: R Interface to X-13ARIMA-SEATS." R package version 1.1.0.

Sax, Christoph, and Peter Steiner. 2013. "Temporal Disaggregation of Time Series." *R Journal*, 5(2): 80–87.

Schmidheiny, Kurt, and Michaela Slotwinski. 2015. "Behavioral Responses to Local Tax Rates: Quasi-Experimental Evidence from a Foreigners Tax Scheme in Switzerland." *CESifo Working Paper*, , (5518).

Schuler, Martin, Pierre Dessemontet, and Dominique Joye. 2005. *Eidgenössische Volkszählung 2000: Die Raumgliederungen der Schweiz.* Swiss Federal Statistical Office.

SECO. 2006. "Technische Notiz. Produktionskonto auf Quartalsbasis und neue BIP-Quartalsreihe." State Secretariat for Economic Affairs.

SECO. 2014*a*. "10. Bericht des Observatoriums zum Freizügigkeitsabkommen Schweiz-EU, Auswirkungen der Personenfreizügigkeit auf den Schweizer Arbeitsmarkt." Staatssekretariat für Wirtschaft (SECO).

SECO. 2014*b*. "Anhang zur Medienmitteilung vom 30. September 2014. Umstellung der vierteljährlichen volkswirtschaftlichen Gesamtrechnungen auf das ESVG 2010." *Press Release Supplemental.*

Sheldon, George. 2007. "Migration, Integration und Wachstum: Die Performance und Wirtschaftliche Auswirkung der Ausländer in der Schweiz." Forschungsstelle für Arbeitsmarkt- und Industrieökonomik (FAI), Universität Basel.

Shih, Kevin, Giovanni Peri, and Chad Sparber. 2015. "STEM Workers, H-1B Visas, and Productivity in US Cities." *Journal of Labor Economics*, 33(S1): 225–255.

Shiskin, Julius, Allan H. Young, and John C. Musgrave. 1965. *The X-11 Variant of the Census Method II Seasonal Adjustment Program.* US Department of Commerce, Bureau of the Census.

Siliverstovs, Boriss. 2011. "The Real-Time Predictive Content of the KOF Economic Barometer." *Schweizerische Zeitschrift fur Volkswirtschaftund Statistik*, 147(3): 353–375.

Siliverstovs, Boriss. 2013. "Dating Business Cycles in Historical Perspective: Evidence for Switzerland." *Jahrbücher für Nationalökonomie & Statistik*, 233(5): 661–679.

Silverman, Bernard W. 1986. *Density Estimation for Statistics and Data Analysis.* London:CRC press.

Speth, Hans-Theo. 2004. "The BV4.1 Procedure for Decomposing and Seasonally Adjusting Economic Time Series." *Methodenberichte*, Heft 3. Statistisches Bundesamt, Wiesbaden.

Spitz-Oener, Alexandra. 2006. "Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure." *Journal of Labor Economics*, 24(2): 235–270.

Stevens, Barrie, Anna di Mattia, and Pierre-Alain Schieb. 2009. "The Future of International Migration." In *The Future of International Migration to OECD Countries.* 17–50. OECD.

Stock, James H., Jonathan H. Wright, and Motohiro Yogo. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business & Economic Statistics*, 20(4): 518–529.

Stone, Richard, David G. Champernowne, and James E. Meade. 1942. "The Precision of National Income Estimates." *The Review of Economic Studies*, 9(2): 111–125.

Swanson, Norman R., and Dick Van Dijk. 2006. "Are Statistical Reporting Agencies Getting it Right? Data Rationality and Business Cycle Asymmetry." *Journal of Business & Economic Statistics*, 24(1): 24–42.

Swiss Federal Statistical Office. 2011. "The New Census." http://www.bfs.admin.ch/bfs/portal/en/index/news/02.html.

Tiller, Richard B., and Thomas D. Evans. 2016. "Methodology for Seasonally Adjusting National Household Survey Labor Force Series with Revisions for 2016." *Current Population Survey Technical Documentation.*

UCDP/PRIO. 2015. "UCDP/PRIO Armed Conflict Dataset v.4-2015." *http://www.pcr.uu.se/research/ucdp/datasets/ucdp_prio_armed_conflict_dataset/.*

U.S. Census Bureau. 2013. "X-13Arima-SEATS Reference Manual." U.S. Census Bureau, Statistical Research Division.

US Department of Labor, Employment and Training Administration. 1977. *Dictionary of Occupational Titles: Fourth Edition.* Washington, DC:US Government Printing Office.

Wright, Jonathan H. 2013. "Unseasonal Seasonals?" *Brookings Papers on Economic Activity*, 47(2): 65–126.

Zellner, Arnold. 1958. "A Statistical Analysis of Provisional Estimates of Gross National Product and its Components, of Selected National Income Components, and of Personal Saving." *Journal of the American Statistical Association*, 53(281): 54–65.